



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 131014

TO: Terra Gibbs
Location: REM/2D10/2C18
Art Unit: 1635
Monday, August 30, 2004

Case Serial Number: 10/029115

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1A69
Phone: 571-272-2518 *BOB*

barbara.obryen@uspto.gov

Search Notes

RUSH

THIS PAGE BLANK (USPTO)

Gibbs, Terra

From: Gibbs, Terra
Sent: Thursday, August 19, 2004 5:21 PM
To: STIC-Biotech/ChemLib
Subject: RE: Sequence comparison

Please modify the search below. The claims are drawn to a nucleic acid at least 98% or 95% identical to SEQ ID NO:1 of USSN 10/029,115.

I need the parameters tweaked such that the sequences are at least 98% identical. If they aren't at least 98% identical, please tweak the parameters such that the sequences are at least 95% identical.

Thank You!

-----Original Message-----

From: Gibbs, Terra
Sent: Thursday, August 19, 2004 5:18 PM
To: STIC-Biotech/ChemLib
Subject: Sequence comparison

1. Could you please compare SEQ ID NO:11 of US 6,656,716 with SEQ ID NO:1 of USSN 10/029,115? I need the parameters tweaked such that the sequences are at least 95% identical.
2. Could you please compare Accession number AB035698 with SEQ ID NO:1 of USSN 10/029,115? I need the parameters tweaked such that the sequences are at least 95% identical.

*Terra Cotta Gibbs, Ph.D.
Art Unit 1635
Remsen Building 2D10
Mailbox 2C18
571-272-0758*

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 28, 2004, 00:43:54, Search time 29 Seconds
(without alignments)
2355.726 Million cell updates/sec

Title: US-10-029-115-2

Perfect score: 6929
Sequence: 1 MCDPAPARSLDDIDLSALRD.....SCGSSQVYFNTLRNCINMW 1312

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_42:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6909	99.7	1332	1	MAK6_HUMAN
2	6464	93.3	1308	1	MAK6_MOUSE
3	4336	62.6	1360	1	TNIX_HUMAN
4	4017.5	58.0	1239	1	MAK4_HUMAN
5	3981.5	57.0	1233	1	MAK4_MOUSE
6	3286	47.4	916	1	TNIX_MOUSE
7	2646	38.2	1080	1	MAK5_MOUSE
8	728	10.7	894	1	MAK3_MOUSE
9	715.5	10.3	862	1	MAK3_RAT
10	694.5	10.0	491	1	STK3_HUMAN
11	693.5	10.0	669	1	HPO_DROME
12	691	10.0	847	1	MAK5_MOUSE
13	687	9.9	846	1	MAK5_HUMAN
14	681.5	9.8	426	1	STK5_HUMAN
15	675.5	9.7	426	1	STK5_MOUSE
16	667	9.6	1080	1	MAK1_YEAST
17	660.5	9.5	833	1	MAK1_HUMAN
18	656.5	9.5	443	1	STK4_HUMAN
19	653.5	9.4	819	1	MAK2_HUMAN
20	644	9.3	968	1	STK4_MOUSE
21	638	9.2	821	1	MAK2_MOUSE
22	636.5	9.2	966	1	MAK3_MOUSE
23	630.5	9.1	691	1	MAK3_MOUSE
24	629	9.1	827	1	NINQ_DROME
25	620.5	9.0	1501	1	STK1_MOUSE
26	552	8.0	471	1	STK1_MOUSE
27	551	8.0	982	1	STK1_MOUSE
28	548	7.9	490	1	STK1_MOUSE
29	548	7.7	658	1	STK1_MOUSE
30	533.5	7.3	971	1	STK1_MOUSE
31	507	7.3	971	1	STK1_MOUSE
32	505	7.3	652	1	STK1_MOUSE
33	504	7.3	556	1	STK1_MOUSE

ALIGNMENTS

RESULT 1			
MAK6_HUMAN	STANDARD:	PRT: 1332 AA.	
ID	OSN4CG: O9P1X1: O9P2R8:		
AC	10-OCT-2003 (Rel. 42, Created)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Mitogen-activated protein kinase kinase kinase 6 (EC 2.7.1.37)		
DE	(MAPK/ERK kinase kinase kinase 6) (MEKKK 6)		
DE	(Mitsugaden/NIK-related kinase) (GCK family kinase MINK)		
GN	MAK6 OR MINK		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Carnivora; Homiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).		
RC	TISSUE=Brain;		
RX	MEDLINE=20175403; PubMed=10708748;		
RA	Dan I., Watanabe N.M., Kobayashi T., Yamashita-Suzuki K., Fukagaya Y.,		
RA	Kajikawa E., Kimura N.K., Nakashima T.M., Matsumoto K.,		
RA	Ninomiya-Tsuji J., Kusumi A.;		
RT	"Molecular cloning of MINK, a novel member of mammalian GCK family		
RT	kinases, which is up-regulated during postnatal mouse cerebral		
RT	development.";		
RT	FEBS Lett. 463:19-23 (2000).		
RN	[2]		
RP	SEQUENCE FROM N.A. (ISOFORM 3).		
RC	TISSUE=Lymph;		
RX	MEDLINE=22386257; PubMed=12477932;		
RA	Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Burow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Matulova K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Starostin M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Udell T.B., Toshiyuki S., Carrinot P., Prange C.,		
RA	Boak S.S., Loggiano N.A., Peters G.J., Abrahamson R.D., Mullaly S.C.,		
RA	Bohak S.A., McMan P.J., McKernan K.J., Malek J.A., Gurrane P.H.,		
RA	Richards S., Kowley D.C., Hale S., Garcia A.H., Gay L.J., Huix S.W.,		
RA	Villalon D.K., Huix D.H., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,		
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,		
RA	Rodriguez Y.S.N., Krzywicki M.I., Skalska U., Smalhus D.E.,		
RA	Schneer A., Schein J.E., Jones S.J.M., Marra M.A.;		
RT	"Generation and initial analysis of more than 15,000 full-length		
RT	human and mouse cDNA sequences.";		
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).		
CC	- FUNCTION: Serine/threonine kinase that may play a role in the		
CC	response to environmental stress. Appears to act upstream of the		
CC	c-Jun N-terminal pathway (by similarity).		
CC	- FUNCTION: May play a role in the development of the brain (by		
CC	similarity).		
CC	- CATALYTIC ACTIVITY: ATP + a protein -> ADP + a		

ALIGNMENTS							
34	502	7.2	553	1	SPAK_RAT	088506	rattus norv
35	500.5	7.2	547	1	SPAK_HUMAN	09uew8	homo sapien
36	498.5	7.2	544	1	PAK3_HUMAN	075914	homo sapien
37	498.5	7.2	544	1	PAK3_RAT	062829	rattus norv
38	496	7.2	544	1	PAK1_RAT	P35465	rattus norv
39	496	7.2	545	1	PAK1_HUMAN	013153	homo sapien
40	495.5	7.2	842	1	CL4_YEAST	P48502	saccharomyc
41	494	7.1	545	1	PAK1_MOUSE	086613	mus musculu
42	493.5	7.1	544	1	PAK3_MOUSE	Q61026	mus musculu
43	488	7.0	589	1	SHK2_MOUSE	Q10056	echinoseach
44	482	7.0	524	1	PAK2_HUMAN	Q13177	homo sapien
45	481.5	6.9	524	1	PAK2_RAT	Q64303	rattus norv

THIS PAGE BLANK (USPTO)

CC phosphoprotein.
CC -1- COPACTOR: Magnesium (By similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=3;
CC IsoId=Q8N4C8-1; Sequence=Displayed;
CC Name=1; Synonyms=MINK-1;
CC IsoId=Q8N4C8-2; Sequence=VSP_007059, VSP_007060;
CC Name=2; Synonyms=MINK-2;
CC IsoId=Q8N4C8-3; Sequence=VSP_007060;
CC -1- TISSUE SPECIFICITY: Expressed in the brain, isoform 2 is more
CC abundant than isoform 1.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC STE20 subfamily.
CC -1- SIMILARITY: Contains 1 CNH domain.
CC -----
CC This SWISS-PROT entry is copyright: It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; AB041926; BAA94838.1; -;
CC ENBL; AB035698; BAA90753.1; -;
CC ENBL; BC034673; AAH34673.1; -;
CC GO; GO:0005524; F:ATP binding; ISS.
CC GO; GO:0004674; F:protein serine/threonine kinase activity; ISS.
CC GO; GO:0007275; P:development; ISS.
CC GO; GO:0006468; P:protein amino acid phosphorylation; ISS.
CC GO; GO:0007243; P:protein kinase cascade; ISS.
CC GO; GO:0006950; P:response to stress; ISS.
CC InterPro: IPR001180; Citron.
CC InterPro: IPR000719; Prot_kinase.
CC InterPro: IPR008271; Ser_thr_pkin_AS.
CC InterPro: IPR002290; Ser_thr_pkinase.
CC Pfam: PF0078C; CNH, 1.
CC Pfam: PF00065; pkinase; 1.
CC ProDom: PD030001; Prot_kinase; 1.
CC SMART: SM00336; CNH, 1.
CC SMART: SM00220; S_TKc; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC Transferrase, Serine/threonine-protein kinase; ATP-binding;
CC Alternative splicing.
CC DOMAIN 25 289 PROTEIN KINASE.
CC DOMAIN 1019 1310 CNH.
CC NF_BIND 31 39 ATP (BY SIMILARITY).
CC BINDING 54 54 ATP (BY SIMILARITY).
CC ACT_SITE 153 153 BY SIMILARITY.
CC VARSPLIC 696 732 Missing (in isoform 1 and isoform
CC 2).
CC FTID=VSP_007059.
CC VARSPLIC 800 800 A -> ASYKFAICE (in isoform 2).
CC FTID=VSP_007060.
CC CONFLICT 771 771 A -> V (IN REF. 1).
CC CONFLICT 775 775 L -> P (IN REF. 1).
CC SEQUENCE 1332 AA; 149809 MW; C80129B652CEB56D CRC64;
CC -----
CC Query Match 99.7%; Score 6909; DB 1; Length 1332;
CC Best Local Similarity 98.5%; Pred. No. 2.9e-202;
CC Matches 1312; Conservative 0; Mismatches 0; Indels 20; Gaps 1;
CC 1 MGDPAARSLDDIDLSALRPAGIFELVEVNGTGYGVYKGRHVKTGOLAAIKWMDVTE 60
CC DB 1 MGDPAARSLDDIDLSALRPAGIFELVEVNGTGYGVYKGRHVKTGOLAAIKWMDVTE 60
CC 61 DEEREIKOENMLKYSHRNIATYVCAITKSPGNDOLWMEFCAGSVTDLVONT 120
CC QY 61 DEEREIKOENMLKYSHRNIATYVCAITKSPGNDOLWMEFCAGSVTDLVONT 120
CC DB 61 DEEREIKOENMLKYSHRNIATYVCAITKSPGNDOLWMEFCAGSVTDLVONT 120

121 KGNALKEDCIAICREILRLGLAHLHAHKVTHRDI KGNVLLTENAELVKLVDFGVSAQLDR 180
121 KGNALKEDCIAICREILRLGLAHLHAHKVTHRDI KGNVLLTENAELVKLVDFGVSAQLDR 180
181 TVGRNTFTIGPYMAEVIACDENPDATYDYSDDTMSLGTATAMAGCAPLCCDHHPWR 240
181 TVGRNTFTIGPYMAEVIACDENPDATYDYSDDTMSLGTATAMAGCAPLCCDHHPWR 240
241 ALFLPRPPRLSKKWSKKFIDFIDTCLIKTYLSRPTEOLLKPPFTRDQTERQVR 300
241 ALFLPRPPRLSKKWSKKFIDFIDTCLIKTYLSRPTEOLLKPPFTRDQTERQVR 300
301 QLNKHDISRKRGRGEETEYVYSGSEEDDSHGEGEFSSIMNVPGESTLREFLRQQ 360
301 QLNKHDISRKRGRGEETEYVYSGSEEDDSHGEGEFSSIMNVPGESTLREFLRQQ 360
361 ENKNSSEALKKQOOLOOQOORDPENAIKHLHCRQRRIEEOKEERRRVEEQORREQRK 420
361 ENKNSSEALKKQOOLOOQOORDPENAIKHLHCRQRRIEEOKEERRRVEEQORREQRK 420
421 LQEKQOORLEDMALREERRQAEEOYKROIEECROSERLORLOQOEHAYLSLQ 480
421 LQEKQOORLEDMALREERRQAEEOYKROIEECROSERLORLOQOEHAYLSLQ 480
481 QOQOQOOLQOQOQOOLLPQDRKPLHYHGRGNPADPKPAWAREVEERTRKMKQONSPLAKS 540
481 QOQOQOOLQOQOQOOLLPQDRKPLHYHGRGNPADPKPAWAREVEERTRKMKQONSPLAKS 540
541 KPGSTGPEPPIPOASPGPPGPLSQTPPMQRPVPEQEGPHK----- 580
541 KPGSTGPEPPIPOASPGPPGPLSQTPPMQRPVPEQEGPHK----- 580
581 SLQDOPTENLAAPFASHDPPDPAIPATATPSARGAVIRONSDPTSEGGPSPNPAWVRP 640
581 SLQDOPTENLAAPFASHDPPDPAIPATATPSARGAVIRONSDPTSEGGPSPNPAWVRP 640
601 DNEAPKVPORSTSIATALNTSCAGGSRPAQAVRAPRNSAWQIYLORRARGT7KPPG 700
601 DNEAPKVPORSTSIATALNTSCAGGSRPAQAVRAPRNSAWQIYLORRARGT7KPPG 700
661 DNEAPKVPORSTSIATALNTSCAGGSRPAQAVRAPRNSAWQIYLORRARGT7KPPG 720
661 DNEAPKVPORSTSIATALNTSCAGGSRPAQAVRAPRNSAWQIYLORRARGT7KPPG 720
701 PPAQPGPPNASSNPDLRRSDPGRWSDSVLPASHGHLPAQGLSLERNVGVASSKLDSSPV 760
701 PPAQPGPPNASSNPDLRRSDPGRWSDSVLPASHGHLPAQGLSLERNVGVASSKLDSSPV 760
721 PPAQPGPPNASSNPDLRRSDPGRWSDSVLPASHGHLPAQGLSLERNVGVASSKLDSSPV 780
721 PPAQPGPPNASSNPDLRRSDPGRWSDSVLPASHGHLPAQGLSLERNVGVASSKLDSSPV 780
761 LSPGHNKAKFDDHRRSRGPRPADFVLLKERTLDDEAPRPPKAMDYSSSEVESSEDDDEEG 820
761 LSPGHNKAKFDDHRRSRGPRPADFVLLKERTLDDEAPRPPKAMDYSSSEVESSEDDDEEG 820
781 LSPGHNKAKFDDHRRSRGPRPADFVLLKERTLDDEAPRPPKAMDYSSSEVESSEDDDEEG 840
781 LSPGHNKAKFDDHRRSRGPRPADFVLLKERTLDDEAPRPPKAMDYSSSEVESSEDDDEEG 840
821 EGCPAEGSRDTPGGRSDGDTDSYSTVWVHDVEITCTOPPYGGTMMVORTPEERENLLH 880
821 EGCPAEGSRDTPGGRSDGDTDSYSTVWVHDVEITCTOPPYGGTMMVORTPEERENLLH 880
841 EGCPAEGSRDTPGGRSDGDTDSYSTVWVHDVEITCTOPPYGGTMMVORTPEERENLLH 900
841 EGCPAEGSRDTPGGRSDGDTDSYSTVWVHDVEITCTOPPYGGTMMVORTPEERENLLH 900
881 ADSNGYTHLPDVVQPSHPTENSKGSPSPKSGSDYQSRGLVKAPGKSSFTPMVLOIY 940
881 ADSNGYTHLPDVVQPSHPTENSKGSPSPKSGSDYQSRGLVKAPGKSSFTPMVLOIY 940
901 ADSNGYTHLPDVVQPSHPTENSKGSPSPKSGSDYQSRGLVKAPGKSSFTPMVLOIY 960
901 ADSNGYTHLPDVVQPSHPTENSKGSPSPKSGSDYQSRGLVKAPGKSSFTPMVLOIY 960
941 QPGSGSDSIPITALVGGEGTRLQDQYDVRKGSVNVNPTNTRAHSETPEIRKYKCRFNS 1000
941 QPGSGSDSIPITALVGGEGTRLQDQYDVRKGSVNVNPTNTRAHSETPEIRKYKCRFNS 1000
961 QPGSGSDSIPITALVGGEGTRLQDQYDVRKGSVNVNPTNTRAHSETPEIRKYKCRFNS 1020
961 QPGSGSDSIPITALVGGEGTRLQDQYDVRKGSVNVNPTNTRAHSETPEIRKYKCRFNS 1020
1001 EILCAALAGVNLVGTENGMLLDRSGQKVGILGRERFQMDVLEGLNLLITISGRN 1060
1001 EILCAALAGVNLVGTENGMLLDRSGQKVGILGRERFQMDVLEGLNLLITISGRN 1060
1021 EILCAALAGVNLVGTENGMLLDRSGQKVGILGRERFQMDVLEGLNLLITISGRN 1080
1021 EILCAALAGVNLVGTENGMLLDRSGQKVGILGRERFQMDVLEGLNLLITISGRN 1080
1061 KLRVYLSLWANKILHRDPEVEKKQGTTVDMEGCGHTRVVKYERIKPLVIALKSSVEV 1120
1061 KLRVYLSLWANKILHRDPEVEKKQGTTVDMEGCGHTRVVKYERIKPLVIALKSSVEV 1120
1081 KLRVYLSLWANKILHRDPEVEKKQGTTVDMEGCGHTRVVKYERIKPLVIALKSSVEV 1140
1081 KLRVYLSLWANKILHRDPEVEKKQGTTVDMEGCGHTRVVKYERIKPLVIALKSSVEV 1140
1121 YAWAPKPTHKMAFKSADLPRLPLVDLTVBEGQRLKVIYGSAGFAHVDVDSGNSYDI 1180
1121 YAWAPKPTHKMAFKSADLPRLPLVDLTVBEGQRLKVIYGSAGFAHVDVDSGNSYDI 1180
1141 YAWAPKPTHKMAFKSADLPRLPLVDLTVBEGQRLKVIYGSAGFAHVDVDSGNSYDI 1200
1141 YAWAPKPTHKMAFKSADLPRLPLVDLTVBEGQRLKVIYGSAGFAHVDVDSGNSYDI 1200
1181 YIPVHIQSIPTPHAIIELENTDGMELLCVEDEGVVNTYGRIIKDVVLQNGEMPTSVAY 1240
1181 YIPVHIQSIPTPHAIIELENTDGMELLCVEDEGVVNTYGRIIKDVVLQNGEMPTSVAY 1240

THIS PAGE BLANK (USPTO)

3 Mon Aug 30 09:54:26 2004

us-10-02

Db 1201 |||||YIPVHIQSQITPHAIIFLPNTDGMEMLLCYEDEGVYVNTYGRITIKDVVLOWGEMPTSVAY 1260
Qy 1241 ICSNQIMGWGEKAIEIRSVETGHLGCVFMKRAQLKFLCERNDKVFFASVRSGGSSQVY 1300
Db 1261 ICSNQIMGWGEKAIEIRSVETGHLGCVFMKRAQLKFLCERNDKVFFASVRSGGSSQVY 1320
Qy 1301 FMTLNRNCIMNW 1312
Db 1321 FMTLNRNCIMNW 1332

THIS PAGE BLANK (uspto)

AUTHORS Dan, I., Watanabe, N.M. and Kusumi, A.

Z

Db 241 AACATCGCCACCTACTACGAGCCTTCATCAAGAGAGAGCCGCCCGGGAAACGATGACCG 300
Qy 307 CTCTGGCTGGTGAATGGAGTTCTGTGGTCTGTGGTTCAGTGAAGTCAAGCTGGTAAAGAACACA 366
Db 301 CTCTGGCTGGTGAATGGAGTTCTGTGGTCTGTGGTTCAGTGAAGTCAAGCTGGTAAAGAACACA 360
Qy 367 AAAGGCAACGCCCTGAAGAGAGACTGTATCGGCTATATCTGACGGAGATCTCTAGGGGT 426
Db 361 AAAGGCAACGCCCTGAAGAGAGACTGTATCGGCTATATCTGACGGAGATCTCTAGGGGT 420
Qy 427 CTGGCCCATCTCATGCCCAACAGGTGATCCATCGAGACATCAAGGGGCGAGAAATGTCTG 486
Db 421 CTGGCCCATCTCATGCCCAACAGGTGATCCATCGAGACATCAAGGGGCGAGAAATGTCTG 480
Qy 487 CTGACAGAGAATGCTGAGGTCAAGCTAGTGGATTTTGGGGTGAAGTCTCAGCTGGACCGC 546
Db 481 CTGACAGAGAATGCTGAGGTCAAGCTAGTGGATTTTGGGGTGAAGTCTCAGCTGGACCGC 540
Qy 547 ACCGTGGGCGAGAGGAACTTTTCATTGGGACTCCCTACTGATGGCTCCAGAGTCAATC 606
Db 541 ACCGTGGGCGAGAGGAACTTTTCATTGGGACTCCCTACTGATGGCTCCAGAGTCAATC 600
Qy 607 GCCTGTGATGAGAACCTCATGACCTATGACCTATGATTAAGAGAGTATATTTGGTCTCTAGA 666
Db 601 GCCTGTGATGAGAACCTCATGACCTATGATTAAGAGAGTATATTTGGTCTCTAGA 660
Qy 667 ATCACAGCCATCGAGATGCGAGAGGAGCCGCCCTCTGTGTGATGACACCCCATGCGA 726
Db 661 ATCACAGCCATCGAGATGCGAGAGGAGCCGCCCTCTGTGTGATGACACCCCATGCGA 720
Qy 727 GGCCTCTTCTCATTTCTCGGAACCTCGGCCAGGCTCAAGTCCAGAGTGTCTAAG 786
Db 721 GGCCTCTTCTCATTTCTCGGAACCTCGGCCAGGCTCAAGTCCAGAGTGTCTAAG 780
Qy 787 AAGTTCAATGACTTCAATGACACATGTCTCATCAAGACTTTACTCGAGCGGCCACCCACG 846
Db 781 AAGTTCAATGACTTCAATGACACATGTCTCATCAAGACTTTACTCGAGCGGCCACCCACG 840
Qy 847 GAGCAGTACTGAAAGTTTCCCTTCAATCGGGACCAAGCCATCGAGCGGCGAGTCCGCATC 906
Db 841 GAGCAGTACTGAAAGTTTCCCTTCAATCGGGACCAAGCCATCGAGCGGCGAGTCCGCATC 900
Qy 907 CAGCTTAAGGACCAATTTGACGATCCCGGAGAGAGCGGGTGAGAAAGAGGAGACAGAA 966
Db 901 CAGCTTAAGGACCAATTTGACGATCCCGGAGAGAGCGGGTGAGAAAGAGGAGACAGAA 960
Qy 967 TATGAGTACAGCGGCGAGGAGGAGAAATGACAGCCATGAGAGAGAAAGAGGACCAAGC 1026
Db 961 TATGAGTACAGCGGCGAGGAGGAGAAATGACAGCCATGAGAGAGAAAGAGGACCAAGC 1020
Qy 1027 TCCATCATGAAAGTGTCTGAGAGTGCATCTTACGCGGGAGTTTCTCGGCTCCAGCAG 1086
Db 1021 TCCATCATGAAAGTGTCTGAGAGTGCATCTTACGCGGGAGTTTCTCGGCTCCAGCAG 1080
Qy 1087 GAAATTAAGAGCACTCAGAGGCTTTAAACAGCAGCAGCAGCTGACGAGCAGCAGCAGCAG 1146
Db 1081 GAAATTAAGAGCACTCAGAGGCTTTAAACAGCAGCAGCAGCTGACGAGCAGCAGCAGCAG 1140
Qy 1147 CGAGACCCCGAGGACACATCAAAACCTGTGCAACAGCGGCGAGCGGCGCATAGAGAG 1206
Db 1141 CGAGACCCCGAGGACACATCAAAACCTGTGCAACAGCGGCGAGCGGCGCATAGAGAG 1200
Qy 1207 CAGAGAGGAGCGGCGCGCTGTGAGAGGAGCAACAGCGGCGGAGCGGAGCAGCGGAAG 1266
Db 1201 CAGAGAGGAGCGGCGCGCTGTGAGAGGAGCAACAGCGGCGGAGCGGAGCAGCGGAAG 1260
Qy 1267 CTGCAGGAGAGCGGCGCGCTGTGAGAGGAGCAATCAAGGCTCTGGCGGGAGGAG 1326
Db 1261 CTGCAGGAGAGAGGAGCAGCGGCGCTGTGAGAGGAGCAATCAAGGCTCTGGCGGGAGGAG 1320
Qy 1327 GAGCGGCGGAGCGGAGCGCTGAGAGGAAATCAAGCGGAAGCAGCTGGAGGAGCAGCGG 1386
Db 1321 GAGCGGCGGAGCGGAGCGCTGAGAGGAAATCAAGCGGAAGCAGCTGGAGGAGCAGCGG 1380

Qy 1387 CAGTCAGAACGTCTTCCAGAGGCACTGCGAGCAGAGCATGCTACCTCAAGTCTCCCTGCGAG 1446
Db 1381 CAGTCAGAACGTCTTCCAGAGGCACTGCGAGCAGAGCATGCTACCTCAAGTCTCCCTGCGAG 1440
Qy 1447 CAGCAGCAACAGCAGCAGCAGCTTCAGAAACAGCAGCAGCAGCAGCTTCCTGCTGGGGAC 1506
Db 1441 CAGCAGCAACAGCAGCAGCAGCTTCAGAAACAGCAGCAGCAGCAGCTTCCTGCTGGGGAC 1500
Qy 1507 AGGAAGCCCTGTATACCATTTATGGTCGGGGCATGAATCCCGCTGACAAACAGACCTTGGGCC 1566
Db 1501 AGGAAGCCCTGTATACCATTTATGGTCGGGGCATGAATCCCGCTGACAAACAGACCTTGGGCC 1560
Qy 1567 CGAGAGGTAGAAGAGAGAAACAGGATGAAACAGCAGCAGCAACTCTCTCTTGGGCCAAGAGC 1626
Db 1561 CGAGAGGTAGAAGAGAGAAACAGGATGAAACAGCAGCAGCAACTCTCTCTTGGGCCAAGAGC 1620
Qy 1627 AAGCCAGGCGACGCGGGCCTGAGCCCCCATCCCCCAGGCTCCCCCAGGGCCCCCAGGA 1686
Db 1621 AAGCCAGGCGACGCGGGCCTGAGCCCCCATCCCCCAGGCTCCCCCAGGGCCCCCAGGA 1680
Qy 1687 CCCCTTTCCAGACTCTCTCTATGACAGGCGCGGTGGAGCCCCCAGGAGGGAACCGACA -- 1744
Db 1681 CCCCTTTCCAGACTCTCTCTATGACAGGCGCGGTGGAGCCCCCAGGAGGGAACCGACAAG 1740
Qy 1745 -----AG 1746
Db 1741 AGCCTGGTGGCACACCGGGTCCCACTGAAGCCATATGACGACCTGTACCCCGATCCCGAG 1800
Qy 1747 TCCCTGAGAGACAGCCACCCGAAACCTGTGCTGCTTCCAGGCTCCCATGACCCCGAC 1806
Db 1801 TCCCTGAGAGACAGCCACCCGAAACCTGTGCTGCTTCCAGGCTCCCATGACCCCGAC 1860
Qy 1807 CCTGCCATCCCGCACCCACTGTGCCACGCCCGAGTGCCTCCGAGAGTGTCTATCCGCGAGAAT 1866
Db 1861 CCTGCCATCCCGCACCCACTGTGCCACGCCCGAGTGCCTCCGAGAGTGTCTATCCGCGAGAAT 1920
Qy 1867 TCAGACCCGACCTTGAAGGACCTTGGGCCCGAGCCCGGAATCCCCAGGCTTGGGTCCGCCCA 1926
Db 1921 TCAGACCCGACCTTGAAGGACCTTGGGCCCGAGCCCGGAATCCCCAGGCTTGGGTCCGCCCA 1980
Qy 1927 GATAAAGAGGCCCCCAGCAAGGTGCTCAGAGGACCTCATCTATCGGCACTGCTCCCTTAAAC 1986
Db 1981 GATAAAGAGGCCCCCAGCAAGGTGCTCAGAGGACCTCATCTATCGGCACTGCTCCCTTAAAC 2040
Qy 1987 ACCAGTGGGCGCGAGGGTCCCGGCCAGCCAGCAGTCCGTGCGCAGACCTTCGACGCAAC 2046
Db 2041 ACCAGTGGGCGCGAGGGTCCCGGCCAGCCAGCAGTCCGTGCGTGC----- 2085
Qy 2047 TCCGCTTGGCAATCTATCTGCAAGGCGGGCAGAGCGGGGCAACCCCAAGCCTCCAGGG 2106
Db 2086 ----- 2085
Qy 2107 CCCCTGCTCAGCCCCCTGGCGCCGCCCAACGCTCTAGTAAACCCCGACCTCAGGAGGAGC 2166
Db 2086 -----AGTAAACCCCGACCTCAGGAGGAGC 2109
Qy 2167 GACCTTGGTGGGAAACGCTCGGACAGCGTCTTTCAGGCTTCTACGCGGCACTTCCCGCAG 2226
Db 2110 GACCTTGGTGGGAAACGCTCGGACAGCGTCTTTCAGGCTTCTACGCGGCACTTCCCGCAG 2169
Qy 2227 GCTGGCTCAGTGGAGCGGAACCGGCTGGGAGCTTCTCCAAACTGGAACAGCTCCCTGTG 2286
Db 2170 GCTGGCTCAGTGGAGCGGAACCGGCTGGGAGTCTCTCCAAACCGGAGAGCTTCCCTGTG 2229
Qy 2287 CTCTCCCTTGGGAATAAGCCCAAGCCCGACGACCAACCGCTCAGCGGCGAGGCGGCGCGCA 2346
Db 2230 CTCTCCCTTGGGAATAAGCCCGACGACCAACCGCTCAGCGGCGAGGCGGCGCGCGCA 2289
Qy 2347 GACTTTGTGTGTAAGAGCGGAGTCTTGGAAGAGGCGGCTTGGGCTTCCCAAGAGAGGCC 2406
Db 2290 GACTTTGTGTGTAAGAGCGGAGTCTTGGAAGAGGCGGCTTGGGCTTCCCAAGAGAGGCC 2349


```
QY 2407 ATGGACTACTCGTCTCCAGCGAGGAGGTGGAAACAGTCAGGACGACGAGGAGGAAGGC 2466
DB |||||
QY 2350 ATGGACTACTCGTCTCCAGCGAGGAGGTGGAAACAGTCAGGACGACGAGGAGGAAGGC 2409
DB |||||
QY 2467 GAAAGCGGGCCACGACAGAGGGAGCAGAGATACCCCTGGGGGCCGCGAGCGATGGGGATACA 2526
DB |||||
QY 2410 GAAAGCGGGCCACGACAGAGGGAGCAGAGATACCCCTGGGGGCCGCGAGCGATGGGGATACA 2469
DB |||||
QY 2527 GACAGCGTCAGCACCATGTGTGTCCAGACGTGCGAGGAGATCACCGGAGACCGACCCCA 2586
DB |||||
QY 2470 GACAGCGTCAGCACCATGTGTGTCCAGACGTGCGAGGAGATCACCGGAGACCGACCCCA 2529
DB |||||
QY 2587 TACGGGGCGGCACCATGTGTGTCCAGAGCGACCCCTGGAAGAGAGCGGAACCTGCTGCAT 2646
DB |||||
QY 2530 TACGGGGCGGCACCATGTGTGTCCAGAGCGACCCCTGGAAGAGAGCGGAACCTGCTGCAT 2589
DB |||||
QY 2647 GCTGACAGCAATGGGGTACACAAACCTGCTGACGTGCTCAGCCCGACGACCTACCCACC 2706
DB |||||
QY 2590 GCTGACAGCAATGGGGTACACAAACCTGCTGACGTGCTCAGCCCGACGACCTACCCACC 2649
DB |||||
QY 2707 GAGAACAGCAAAAGGCCAAAGCCCAACCTCGAAGGATGGGAGTGGTGACTACCACTCTCGT 2766
DB |||||
QY 2650 GAGAACAGCAAAAGGCCAAAGCCCAACCTCGAAGGATGGGAGTGGTGACTACCACTCTCGT 2709
DB |||||
QY 2767 GGGCTGGTAAAGCCCTGGCAGAGAGCTCGTTACAGATGTTTGTGATCTAGGGATCTAC 2826
DB |||||
QY 2710 GGGCTGGTAAAGCCCTGGCAGAGAGCTCGTTACAGATGTTTGTGATCTAGGGATCTAC 2769
DB |||||
QY 2827 CAGCCTGGAGGACATGGGGACAGACATCCCCATCACAGCCCTAGTGGTGGAGAGGGCACT 2886
DB |||||
QY 2770 CAGCCTGGAGGACATGGGGACAGACATCCCCATCACAGCCCTAGTGGTGGAGAGGGCACT 2829
DB |||||
QY 2887 CGGCTCGACCACTGTCAGTACGACGTGAGGAAGGGTCTGTGGTCAACGTGAATCCCAAC 2946
DB |||||
QY 2830 CGGCTCGACCACTGTCAGTACGACGTGAGGAAGGGTCTGTGGTCAACGTGAATCCCAAC 2889
DB |||||
QY 2947 AACACCGGGCCACAGTCAGACCCCTGAGATCCCGAAGTACAGAAGCGATTCAACTCC 3006
DB |||||
QY 2890 AACACCGGGCCACAGTCAGACCCCTGAGATCCCGAAGTACAGAAGCGATTCAACTCC 2949
DB |||||
QY 3007 GAGATCCTCTGTGCAGCCCTTTGGGGGGTCAACCTGCTGGTGGGCAACCGGGCTG 3066
DB |||||
QY 2950 GAGATCCTCTGTGCAGCCCTTTGGGGGGTCAACCTGCTGGTGGGCAACCGGGCTG 3009
DB |||||
QY 3067 ATGTTGCTGGAACGAGTGGGCAAGGGATGTATGGATCACTATGGGGCGGACGCTTC 3126
DB |||||
QY 3010 ATGTTGCTGGAACGAGTGGGCAAGGGATGTATGGATCACTATGGGGCGGACGCTTC 3069
DB |||||
QY 3127 CAGCAGATGGATGTGTGGAGGGGCTCAACCTGCTCATCACCATCTCAGGGAAAGGAAC 3186
DB |||||
QY 3070 CAGCAGATGGATGTGTGGAGGGGCTCAACCTGCTCATCACCATCTCAGGGAAAGGAAC 3129
DB |||||
QY 3187 AAACCTGGGGTGTATTACCTGTCTTGGCTCCGGAAACAGATTCTGCACAATGACCCAGAA 3246
DB |||||
QY 3130 AAACCTGGGGTGTATTACCTGTCTTGGCTCCGGAAACAGATTCTGCACAATGACCCAGAA 3189
DB |||||
QY 3247 GTGGAGAAGAACGAGGGCTGGAACACCGTGGGGGGAATGAGAGGGCTGCGGGGCACTACCGT 3306
DB |||||
QY 3190 GTGGAGAAGAACGAGGGCTGGAACACCGTGGGGGGAATGAGAGGGCTGCGGGGCACTACCGT 3249
DB |||||
QY 3307 GTTGTGAATACAGCGGATTAAGTTTCCCTGGTTCATCGCCCTCAAGAGCTCCGTTGAGAGTG 3366
DB |||||
QY 3250 GTTGTGAATACAGCGGATTAAGTTTCCCTGGTTCATCGCCCTCAAGAGCTCCGTTGAGAGTG 3309
DB |||||
QY 3367 TATGCTGGGCCCCCAACCTTACCAAAATTCATGGCCCTTCAAGTCTCTTTGCCGACCTC 3426
DB |||||
QY 3310 TATGCTGGGCCCCCAACCTTACCAAAATTCATGGCCCTTCAAGTCTCTTTGCCGACCTC 3369
DB |||||
QY 3427 CCCACCGCCCTCTGTGTGTGACCTTGAACAGTGAAGAGGGGAGCGGCTCAAGTCAATC 3486
DB |||||
QY 3370 CCCACCGCCCTCTGTGTGTGACCTTGAACAGTGAAGAGGGGAGCGGCTCAAGTCAATC 3429
DB |||||
QY 3487 TATGGCTCCAGTGTGCTTCCATGTGTGTGATGAGAACCCCTGTATGCCACCTATGATTTACAGAGTG 3546
DB |||||
```

```
DB 3430 TATGGCTCCAGTGTGCTTCCATGTGTGTGATGTGCACTCGGGGAAACAGCTATGACATC 3489
QY |||||
DB 3547 TACATCCCTGTGCATCCAGAGCCAGATCAGCCCCCATGCATCATCTTCTCCCTCCCAAC 3606
DB |||||
DB 3490 TACATCCCTGTGCATCCAGAGCCAGATCAGCCCCCATGCATCATCTTCTCCCTCCCAAC 3549
QY |||||
DB 3607 ACCGACGGCATGGAGATGCTGTGTCTACGAGACGAGGGTGTCTACGTCAACACGTAC 3666
DB |||||
DB 3550 ACCGACGGCATGGAGATGCTGTGTCTACGAGACGAGGGTGTCTACGTCAACACGTAC 3609
QY |||||
DB 3667 GGGCGCATCATTAAGGATGTGTGTCTCCAGTGGGGGAGATGCTTCTTGTGGGCTAC 3726
DB |||||
DB 3610 GGGCGCATCATTAAGGATGTGTGTCTCCAGTGGGGGAGATGCTTCTTGTGGGCTAC 3669
QY |||||
DB 3727 ATCTGCTCCAACCAAGATAATGGGCTGGGGTGCAGAAAGCCATTGAGATCCGCTCTGTGTGAG 3786
DB |||||
DB 3670 ATCTGCTCCAACCAAGATAATGGGCTGGGGTGCAGAAAGCCATTGAGATCCGCTCTGTGTGAG 3729
QY |||||
DB 3787 ACGGGCCACCTCGACGGGGTCTTTCATGCAACAAACAGCTCAGAGGCTCAAGTCTCTGTGT 3846
DB |||||
DB 3730 ACGGGCCACCTCGACGGGGTCTTTCATGCAACAAACAGCTCAGAGGCTCAAGTCTCTGTGT 3789
QY |||||
DB 3847 GAGCGGAATGACAAAGGTGTTTTTGGCCTCAGTCCGCTCTGGGGGAGAGCAAGTTTAC 3906
DB |||||
DB 3790 GAGCGGAATGACAAAGGTGTTTTTGGCCTCAGTCCGCTCTGGGGGAGAGCAAGTTTAC 3849
QY |||||
DB 3907 TTTCATGACTCTGAACCGTAACTGCATCATGAACCTGGTGA 3945
DB |||||
DB 3850 TTTCATGACTCTGAACCGTAACTGCATCATGAACCTGGTGA 3888
DB |||||
```

RESULT 2
us-09-688-188b-11

Query Match 95.8%; Score 3785; DB 2; Length 4133;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 3824; Conservative 0; Mismatches 15; Indels 150; Gaps 3;

```
QY 110 CTTAGCGACAGGTGTACAAAGGTCGGCATGTCAAGAGCGGCGAGCTGGCTGCATCAAGG 169
DB 2 CATTGGGGAGGGTGTATAGGGTCTGGCATGTCAAGAGCGGCGAGCTGGCTGCATCAAGG 61
QY 170 TCATGATGTCAACGAGGACGAGGAGGAAGAGATCAAAACAGGAGATCAACATGCTGAAAA 229
DB 62 TCATGATGTCAACGAGGACGAGGAGGAAGAGATCAAAACAGGAGATCAACATGCTGAAAA 121
QY 230 AGTACTCTCAACACCGCAACATCGCCACCTACTACGAGCCCTTCAAGAGAGAGCCCCC 289
DB 122 AGTACTCTCAACACCGCAACATCGCCACCTACTACGAGCCCTTCAAGAGAGAGCCCCC 181
QY 290 CGGGAACAGATGACACGAGCTCTGGCTGGTGTGAGTGGAGTTCCTGTGGTCTGTTTCAGTGA 349
DB 182 CGGGAACAGATGACACGAGCTCTGGCTGGTGTGAGTGGAGTTCCTGTGGTCTGTTTCAGTGA 241
QY 350 ACCTGGTAAAGAACACAAAAGGCCAACCGCCCTGAAGGAGGACTGTATCGCCTATATCTGCA 409
DB 242 ACCTGGTAAAGAACACAAAAGGCCAACCGCCCTGAAGGAGGACTGTATCGCCTATATCTGCA 301
QY 410 GGGAGATCCTCAGGGGTCTGGCCCATCTCCATGCCCCAACAGGTGATCCATCGAGACATCA 469
DB 302 GGGAGATCCTCAGGGGTCTGGCCCATCTCCATGCCCCAACAGGTGATCCATCGAGACATCA 361
QY 470 AGGGGACAGATGTGCTGTGTCAGAGAAATGCTGAGGTCAAGTGTGAGTGGGTTTGGGGTGA 529
DB 362 AGGGGACAGATGTGCTGTGTCAGAGAAATGCTGAGGTCAAGTGTGAGTGGGTTTGGGGTGA 421
QY 530 GTGCTCAGCTGACCGCACCGCTGGGCGAGACGGAACACTTTTCAATTGGGACTCCCTACTGGA 589
DB 422 GTGCTCAGCTGACCGCACCGCTGGGCGAGACGGAACACTTTTCAATTGGGACTCCCTACTGGA 481
QY 590 TGGCTCCAGAGGTCAATCGCCTGTGTGATGAGAACCCCTGTATGCCACCTATGATTTACAGAGTG 649
DB |||||
```

Db 482 TGGCTCCAGAGTCAATCGCTGTGATGAGAACCCCTGATGCCACCTATGATTACAGGAGTG 541
Qy 650 ATATTGTGTTCTTAGGAATACAGCCATCGAGATGGCAGAGGGAGCCCCCTCTGTGTG 709
Db 542 ATATTGTGTTCTTAGGAATACAGCCATCGAGATGGCAGAGGGAGCCCCCTCTGTGTG 601
Qy 710 ACATGACCCCATCGAGGCCCTTCTCTCATCTTCGGAAACCTTCGGCCGACGGTCAAGT 769
Db 602 ACATGACCCCATCGAGGCCCTTCTCTCATCTTCGGAAACCTTCGGCCGACGGTCAAGT 661
Qy 770 CCAAGAAGTGTCTAAGAAGTTTCAATTGACTTCATTGACACATGTCTCATCAAGACTTACC 829
Db 662 CCAAGAAGTGTCTAAGAAGTTTCAATTGACTTCATTGACACATGTCTCATCAAGACTTACC 721
Qy 830 TGAGCGCCGCCACCCACGGAGCAGCTACTGAAGTTTCCCTTCATCCGGACACGAGCCACGG 889
Db 722 TGAGCGCCGCCACCCACGGAGCAGCTACTGAAGTTTCCCTTCATCCGGACACGAGCCACGG 781
Qy 890 AGCGCAGGTCCGCATCCAGCTTAAGGACCAATTGACCGATCCCGGAAGAGCGGGTG 949
Db 782 AGCGCAGGTCCGCATCCAGCTTAAGGACCAATTGACCGATCCCGGAAGAGCGGGTG 841
Qy 950 AGAAAGAGGAGACAGAAATATAGTACAGCGGAGGAGGAGGAGATGACAGCCATGGAG 1009
Db 842 AGAAAGAGGAGACAGAAATATAGTACAGCGGAGGAGGAGGAGATGACAGCCATGGAG 901
Qy 1010 AGAAAGGAGAGCAAGCTCCATCATGAACGTGCTGGAGAGTCGACTTACCGCGGGAGT 1069
Db 902 AGAAAGGAGAGCAAGCTCCATCATGAACGTGCTGGAGAGTCGACTTACCGCGGGAGT 961
Qy 1070 TTCTCGGCTCAGCAGGAAATAGAGCACTCAGAGGCTTTAAACAGCAGCAGCAGC 1129
Db 962 TTCTCGGCTCAGCAGGAAATAGAGCACTCAGAGGCTTTAAACAGCAGCAGCAGC 1021
Qy 1130 TGCAGCAGCAGCAGCAGAGACCCCGAGGACACATCAAAACCTGCTGCAACAGCGGC 1189
Db 1022 TGCAGCAGCAGCAGCAGAGACCCCGAGGACACATCAAAACCTGCTGCAACAGCGGC 1081
Qy 1190 AGCGCGCATAGAGAGCAGAAAGAGAGCGCGCGCTGGAGGAGCAACAGCGCGCGG 1249
Db 1082 AGCGCGCATAGAGAGCAGAAAGAGAGCGCGCGCTGGAGGAGCAACAGCGCGCGG 1141
Qy 1250 AGCGGAGCAGCGGAAGCTGCAGGAGAGGAGCAGCGCGCTGGAGGACATGCAAG 1309
Db 1142 AGCGGAGCAGCGGAAGCTGCAGGAGAGGAGCAGCGCGCTGGAGGACATGCAAG 1201
Qy 1310 CTCTCGGCGGAGGAGAGCGCGGCGCAGCGGAGCGTGCAGCA ----- 1352
Db 1202 CTCTCGGCGGAGGAGAGCGCGGCGCAGCGGAGCGGAGCGGAGCAATATTGTCACA 1261
Qy 1353 ----- 1352
Db 1262 GGCTAGAGGAGGAGCAGCAGCTCGAGATCCTTTCAGCAACAGCTGCTCCAGGAAACAGG 1321
Qy 1353 -----GGAATACAGCGGAGCAGCTGGAGGAGCAGCGCAGTCAGAACTCTCC 1402
Db 1322 CCCTGCTGTGTAATACAAAGCGGAAGCAGCTGGAGGAGCGCGCAGTCAGAACTCTCC 1381
Qy 1403 AGAGGCAGCTGCAGCAGGAGCAGTACCTCAAGTCCCTGCAGCAGCAGCAACAGCAGC 1462
Db 1382 AGAGGCAGCTGCAGCAGGAGCATGCTACCTCAAGTCCCTGCAGCAGCAGCAACAGCAGC 1441
Qy 1463 AGCAGCTTCAGAAAACAGCAGCAGCAGCAGTCTCTGCTTGGGGAACAGGAAGCCCTGTACC 1522
Db 1442 AGCAGCTTCAGAAAACAGCAGCAGCAGCAGTCTCTGCTTGGGGAACAGGAAGCCCTGTACC 1501
Qy 1523 ATTATGCTCGGGCATGAATCCCGCTGCAAAACAGCAGCTGGGCCGAGAGGTAGAAAGA 1582
Db 1502 ATTATGCTCGGGCATGAATCCCGCTGCAAAACAGCAGCTGGGCCGAGAGGTAGAAAGA 1561
Qy 1583 GAAACAGGATGAAACAGCAGCAGCAACTCTTCCCTTGGCCAAAGCAGCAGCAGCAGCAG 1642
Db 1562 GAAACAGGATGAAACAGCAGCAGCAACTCTTCCCTTGGCCAAAGCAGCAGCAGCAGCAG 1621

Qy 1643 GGCTTGAGCCCCCATATCCCCAGGCTTCCCAAGGGCCCCCAGAGACCCCTTTCCAGACTC 1702
Db 1622 GGCTTGAGCCCCCATATCCCCAGGCTTCCCAAGGGCCCCCAGAGACCCCTTTCCAGACTC 1681
Qy 1703 CTCCTATGCAAGGCGGCTGAGCCCCCAGAGAGGAGACCGACACA ----- 1744
Db 1682 CTCCTATGCAAGGCGGCTGAGCCCCCAGAGGAGACCGACACAAGAGAGCTGTGTGGCACACC 1741
Qy 1745 -----AGTCCCTGCAGGACCGC 1762
Db 1742 GGGTCCCACTGAAGCCATATGCAGCACCTGTATCCCGATCCCAGTCCCTGCAGGACCGC 1801
Qy 1763 CCACCCGAAATCTGGCTGCTTCCAGGCTCCCATGACCCCGACCTTGCATCCCCGAC 1822
Db 1802 CCACCCGAAATCTGGCTGCTTCCAGGCTCCCATGACCCCGACCTTGCATCCCCGAC 1861
Qy 1823 CCATGCGACCGCCAGTCCCGAGGAGCTGTATCCGCCAGAAATTCAGACCCCACTCTG 1882
Db 1862 CCATGCGACCGCCAGTCCCGAGGAGCTGTATCCGCCAGAAATTCAGACCCCACTCTG 1921
Qy 1883 AAGGACCTGGCCCCCAGCCCGAATCCCCAGCTTGGGTCCGCCAGATAACGAGGCCCCAC 1942
Db 1922 AAGGACCTGGCCCCCAGCCCGAATCCCCAGCTTGGGTCCGCCAGATAACGAGGCCCCAC 1981
Qy 1943 CCAAGGTGCTCAGAGGACCTCATCTATCGCCACTGCCCCCTTAACACAGTGGGGCCGAG 2002
Db 1982 CCAAGGTGCTCAGAGGACCTCATCTATCGCCACTGCCCCCTTAACACAGTGGGGCCGAG 2041
Qy 2003 GGTCCCGCCAGCCAGGAGCTCGTGCCAGAGCTCGCAGCAACTCCGCTGGCAAAATCT 2062
Db 2042 GGTCCCGCCAGCCAGGAGCTCGTGCCAGAGCTCGCAGCAACTCCGCTGGCAAAATCT 2101
Qy 2063 ATCTGCAAAAGCGCGGACAGCGGGCACCCCAAGCCTCCAGGGCCCCCTTGTCTCAGCCCC 2122
Db 2102 ATCTGCAAAAGCGCGGACAGCGGGCACCCCAAGCCTCCAGGGCCCCCTTGTCTCAGCCCC 2161
Qy 2123 CTGGCCCGCCCAAGCGCTCTAGTAAACCCGACCTCAGGAGGAGCGACCTGGCTGGGAAC 2182
Db 2162 CTGGCCCGCCCAAGCGCTCTAGTAAACCCGACCTCAGGAGGAGCGACCTGGCTGGGAAC 2221
Qy 2183 GCTCGGACAGCTCTTTCAGGCTCTCAAGGGACCTTCCCGCAGGCTGGCTCACTGGAGC 2242
Db 2222 GCTCGGACAGGCTCTTTCAGGCTCTCAAGGGACCTTCCCGCAGGCTGGCTCACTGGAGC 2281
Qy 2243 GGAACCGCTGGGAGCTCTTCCAAACTGACAGTCCCTGTGCTCTCCCTTGGGAATA 2302
Db 2282 GGAACCGCTGGGAGTCTCTTCCAAACCGGACAGCTCCCTGTGCTCTCCCTTGGGAATA 2341
Qy 2303 AAGCCAAAGCCGAGCACCGCTCAGCGCCAGCGCCGCGCCGCGAGACTTTGTGTGCTGA 2362
Db 2342 AAGCCAAAGCCGAGCACCGCTCAGCGCCAGCGCCGCGCCGCGAGACTTTGTGTGCTGA 2401
Qy 2363 AAGAGCGGACTCTGGAGCAGGCCCCCTCGGCTCTCCCAAGAGGGCCATGGACTACTCTCGT 2422
Db 2402 AAGAGCGGACTCTGGAGCAGGCCCCCTCGGCTCTCCCAAGAGGGCCATGGACTACTCTCGT 2461
Qy 2423 CCAGCAGGAGGTGGAAGCAGTGAGGACCAAGAGGAGGAAGGCAAGGGCGGCGCAGCAG 2482
Db 2462 CCAGCAGGAGGTGGAAGCAGTGAGGACCAAGAGGAGGAAGGCAAGGGCGGCGCAGCAG 2521
Qy 2483 AGGGAGCAGAGATACCCCTGGGGCCGCGAGCGATGGGGATACAGACAGCGTTCAGACCA 2542
Db 2522 AGGGAGCAGAGATACCCCTGGGGCCGCGAGCGATACAGACAGCGTTCAGACCA 2578
Qy 2543 TGTGTGTCCAGCAGCTCAGGAGATCACCGGGACCCAGCCCCCATATCGGGGGCGGCA 2602
Db 2579 TGTGTGTCCAGCAGCTCAGGAGATCACCGGGACCCAGCCCCCATATCGGGGGCGGCA 2638
Qy 2603 TGTGTGTCCAGCAGCAGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2662
Db 2639 TGTGTGTCCAGCAGCAGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2698

QY	2663	ACACAAACCTGCTGACGTGGTCCAGCCAGCACTCACCCACGAGAACAGAAAGGCC	2722
Db	2699	ACACAAACCTGCTGACGTGGTCCAGCCAGCACTCACCCACGAGAACAGAAAGGCC	2758
QY	2723	AAAGCCACCCCTCGAAGGATGGAGTGTGACTACCACTGCTGCTGGGTGTGTAAGGCC	2782
Db	2759	AAAGCCACCCCTCGAAGGATGGAGTGTGACTACCACTGCTGCTGGGTGTGTAAGGCC	2818
QY	2783	CTGGCAGAGCTCGTTCACGATGTTGTGGATCTAGGGATCTTACAGAGCTCGAGGACGTG	2842
Db	2819	CTGGCAGAGCTCGTTCACGATGTTGTGGATCTAGGGATCTTACAGAGCTCGAGGACGTG	2878
QY	2843	GGCAGACGATCCCAATCACAGCCCTAGTGGGTGGAGAGGCACTCGGCTCCAGACAGCTGC	2902
Db	2879	GGGACAGCATCCCAATCACAGCCCTAGTGGGTGGAGAGGCACTCGGCTCCAGACAGCTGC	2938
QY	2903	AGTACGACGTGAGGAAGGTTCTGTGGTCAACGTTGAAATCCCAACCAACCCGGGCCACAC	2962
Db	2939	AGTACGACGTGAGGAAGGTTCTGTGGTCAACGTTGAAATCCCAACCAACCCGGGCCACAC	2998
QY	2963	GTGAGACCCCTGAGATCCGGAAGTACAGAGCGATTCAACTCCGAGATCCCTCTGTGCAG	3022
Db	2999	GTGAGACCCCTGAGATCCGGAAGTACAGAGCGATTCAACTCCGAGATCCCTCTGTGCAG	3058
QY	3023	CCCTTTGGGGGCTCAACCTGCTGGTGGGCA CGGAGAACGGGCTGATGTTGCTGGACCGAA	3082
Db	3059	CCCTTTGGGGGCTCAACCTGCTGGTGGGCA CGGAGAACGGGCTGATGTTGCTGGACCGAA	3118
QY	3083	GTGGGACGGGCAAGGTGTATGGACTCATTTGGGCGGCGACGTTCCAGCAGATGGATGTGC	3142
Db	3119	GTGGGACGGGCAAGGTGTATGGACTCATTTGGGCGGCGACGTTCCAGCAGATGGATGTGC	3178
QY	3143	TGGAGGGGCTCAACCTGCTCATCAACCTCAGGGAAGGAAACAACTGCGGGGTGATT	3202
Db	3179	TGGAGGGGCTCAACCTGCTCATCAACCTCAGGGAAGGAAACAACTGCGGGGTGATT	3238
QY	3203	ACCTGTCTGGGCTCCGGAACAAGATTCTGCACATGACCAAGAGTGGAGAAAGAGCAGG	3262
Db	3239	ACTTGTCTGGGCTCCGGAACAAGATTCTGCACATGACCAAGAGTGGAGAAAGAGCAGG	3298
QY	3263	GCTGGACCA CCGTGGGGGACATGGAGGGCTGCGGGCACTACCGTGTGTGAAATACGAGC	3322
Db	3299	GCTGGACCA CCGTGGGGGACATGGAGGGCTGCGGGCACTACCGTGTGTGAAATACGAGC	3358
QY	3323	GGATTAAAGTTCTGTGTCATCGCCCTCAAGAGCTCCGTTGAGGTGTATGCTGGGGCCCCA	3382
Db	3359	GGATTAAAGTTCTGTGTCATCGCCCTCAAGAGCTCCGTTGAGGTGTATGCTGGGGCCCCA	3418
QY	3383	AACCCCTACCAAAATTCATGGGCTTCAAGTCCCTTTCGCGACCTCCGCCACCGCCCTCTGC	3442
Db	3419	AACCCCTACCAAAATTCATGGGCTTCAAGTCCCTTTCGCGACCTCCGCCACCGCCCTCTGC	3478
QY	3443	TGGTCGACCTGACAGTAGAGAGGGGACGCGCTCAAGGTCACTATATGCTCCAGTGTG	3502
Db	3479	TGGTCGACCTGACAGTAGAGAGGGGACGCGCTCAAGGTCACTATATGCTCCAGTGTG	3538
QY	3503	GCTTCCATGCTGTGGATGTGACTCGGGGACAGCTATGACATCTACATCCCTGTGCACA	3562
Db	3539	GCTTCCATGCTGTGGATGTGACTCGGGGACAGCTATGACATCTACATCCCTGTGCACA	3598
QY	3563	TCGAGAGCCAGATCACGCCCATGCCATCATCTTCCTCCCAACACCGGCGCATGGAGA	3622
Db	3599	TCGAGAGCCAGATCACGCCCATGCCATCATCTTCCTCCCAACACCGGCGCATGGAGA	3658
QY	3623	TGCTGCTGTGCTACGAGGACGAGGGTGTCTACGTCAACACGTAACGGGCGCATCAATTAAGG	3682
Db	3659	TGCTGCTGTGCTACGAGGACGAGGGTGTCTACGTCAACACGTAACGGGCGCATCAATTAAGG	3718
QY	3683	ATGTGTGCTGACAGTGGGGGAGATGCGCTACTTCTGTGGGCTACATCTGCTCCAACCGA	3742
Db	3719	ATGTGTGCTGACAGTGGGGGAGATGCGCTACTTCTGTGGGCTACATCTGCTCCAACCGA	3778
QY	3743	TAAATGGGCTGGGGTGAGAAAGCCATTGAGATCCGCTCTGTGGAGACGGGCCACCTCGACG	3802

Db	3779	TAATGGGCTGGGGTGAGAAAGCCATTGAGATCCGCTCTGTGGACCGGCCACCTCGACG	3838
QY	3803	GGGTCTTCATGCAACAACGAGCTCAGAGGCTCAAGTTCTGTGTGAGCGGAATGACAAGG	3862
Db	3839	GGGTCTTCATGCAACAACGAGCTCAGAGGCTCAAGTTCTGTGTGAGCGGAATGACAAGG	3898
QY	3863	TGTTTTTTTGGCTCAGTCCGCTCTGCGGGGCGAGCAGCCAAAGTTTACTTATGACTCTGAAC	3922
Db	3899	TGTTTTTTTGGCTCAGTCCGCTCTGCGGGGCGAGCAGCCAAAGTTTACTTATGACTCTGAAC	3958
QY	3923	GTAACCTGATCATGAACCTGGTGAAGGGC	3951
Db	3959	GTAACCTGATCATGAACCTGGTGAAGGGC	3987

Search completed: August 26, 2004, 15:39:57
Job time : 23 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2004, 00:43:54 ; Search time 29 Seconds
(without alignments)
2355.726 Million cell updates/sec

Title: US-10-029-115-2
Perfect score: 6929
Sequence: 1 MGDPAFARSLDIDLSALRD.....SGSSQVFTLNRCIMNW 1312

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6909	99.7	1332	1	Q84C8 homo sapien
2	6764	93.3	1308	1	Q9J52 mus musculus
3	4336	62.6	1360	1	Q9UK5 homo sapien
4	4017.5	58.0	1239	1	Q95819 homo sapien
5	3951.5	57.0	1233	1	P97820 mus musculus
6	3286	47.4	916	1	P83510 mus musculus
7	2646	38.2	1080	1	Q23356 caenorhabdi
8	738	10.7	894	1	Q81V8 homo sapien
9	729	10.5	862	1	Q92412 rattus norv
10	715.5	10.3	487	1	Q13043 homo sapien
11	694.5	10.0	491	1	Q13188 homo sapien
12	693.5	10.0	669	1	Q8T086 drosophila
13	691	10.0	847	1	Q8BPM2 mus musculus
14	687	9.9	846	1	Q9V4K4 homo sapien
15	681.5	9.8	426	1	Q00506 homo sapien
16	675.5	9.7	426	1	Q922W1 mus musculus
17	667	9.6	1080	1	P38692 saccharomyc
18	660.5	9.5	833	1	Q92918 homo sapien
19	656.5	9.5	443	1	Q9V6E0 homo sapien
20	653.5	9.4	819	1	Q12851 h mitogen-a
21	644	9.3	968	1	Q94804 homo sapien
22	638	9.2	821	1	Q61161 mus musculus
23	636.5	9.2	966	1	O55098 mus musculus
24	630.5	9.1	681	1	Q99JP0 mus musculus
25	629	9.1	827	1	P70218 mus musculus
26	620.5	9.0	1501	1	P10676 drosophila
27	552	8.0	471	1	O14305 schizosacch
28	551	8.0	982	1	P46549 caenorhabdi
29	548	7.9	490	1	P08458 saccharomyc
30	533.5	7.7	658	1	P50527 schizosacch
31	507	7.3	971	1	O14427 candida alb
32	505	7.3	652	1	O75011 schizosacch
33	504	7.3	556	1	Q921W9 mus musculus

ALIGNMENTS

RESULT 1

34	502	7.2	553	1	SPAK_RAT	O88506 rattus norv
35	500.5	7.2	547	1	SPAK_HUMAN	Q9UEW8 homo sapien
36	498.5	7.2	544	1	PAK3_HUMAN	O75914 homo sapien
37	498.5	7.2	544	1	PAK3_RAT	Q62829 rattus norv
38	496	7.2	544	1	PAK1_RAT	P35465 rattus norv
39	496	7.2	545	1	PAK1_HUMAN	Q13153 homo sapien
40	495.5	7.2	842	1	CLA4_YEAST	P48562 saccharomyc
41	494	7.1	545	1	PAK1_MOUSE	O88643 mus musculus
42	493.5	7.1	544	1	PAK3_MOUSE	Q61036 mus musculus
43	488	7.0	589	1	SHK2_SCHPO	Q10056 schizosacch
44	482	7.0	524	1	PAK2_HUMAN	Q13177 homo sapien
45	481.5	6.9	524	1	PAK2_RAT	Q64303 rattus norv
ID	MAK6_HUMAN	STANDARD;	PRT;	1332	AA.	
AC	Q8N4C8; Q9P1X1; Q9P2R8;					
DT	10-OCT-2003 (Rel. 42, Created)					
DT	10-OCT-2003 (Rel. 42, Last sequence update)					
DT	10-OCT-2003 (Rel. 42, Last annotation update)					
DE	Mitogen-activated protein kinase kinase kinase 6 (EC 2.7.1.37)					
DE	(MAPK/ERK kinase kinase kinase 6) (MEK kinase kinase 6) (MEKK 6)					
DE	(Misshapen/NIK-related kinase) (GCK family kinase MINK).					
GN	MAP4K6 OR MINK.					
OS	Homo sapiens (Human)					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
OX	NCBI_TaxID=9606;					
RN	[1]					
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).					
RC	TISSUE=Brain;					
RX	MEDLINE=20175403; PubMed=10708748;					
RA	Dan I., Watanabe N.M., Kobayashi T., Yamashita-Suzuki K., Fukagaya Y.,					
RA	Kajikawa E., Kimura W.K., Nakashima T.M., Matsumoto K.,					
RA	Ninomiya-Tsuji J., Kusumi A.;					
RT	"Molecular cloning of MINK, a novel member of mammalian GCK family					
RT	kinases, which is up-regulated during postnatal mouse cerebral					
RT	development.";					
FEBS	Letts. 469:19-23 (2000).					
[2]						
RP	SEQUENCE FROM N.A. (ISOFORM 3).					
RC	TISSUE=Lymph;					
RA	MEDLINE=22388257; PubMed=12477932;					
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,					
RA	Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,					
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,					
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,					
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,					
RA	Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,					
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,					
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,					
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,					
RA	Villalobos D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,					
RA	Hahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,					
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,					
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,					
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,					
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,					
RA	Schneerch A., Schein J.B., Jones S.J.M., Marra M.A.;					
RT	"Generation and initial analysis of more than 15,000 full-length					
RT	human and mouse cDNA sequences.";					
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).					
CC	-!- FUNCTION: Serine/threonine kinase that may play a role in the					
CC	response to environmental stress. Appears to act upstream of the					
CC	C-jun N-terminal pathway (By similarity).					
CC	-!- FUNCTION: May play a role in the development of the brain (By					
CC	similarity)					
CC	-!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a					

Db 1201 YIPVHIQSIITPHALIFIPNTDGMELLCYEDGVYVNTYGRILIKDVVLQWGMETSVA 1260
 1241 ICNSQIMGMGEKAIEIRSVETGHLGDFVPMHKAQRLKFLCERNDRKVFASVRSGGSSQVY 1300
 1261 ICNSQIMGMGEKAIEIRSVETGHLGDFVPMHKAQRLKFLCERNDRKVFASVRSGGSSQVY 1320
 QY 1301 FMTLNRNCIMNW 1312
 Db 1321 FMTLNRNCIMNW 1332

RESULT 2

MAK6 MOUSE
 ID MAK6 MOUSE STANDARD; PRT; 1308 AA.
 AC Q9JW52; Q921MG; Q9JW92;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Mitogen-activated protein kinase kinase kinase 6 (EC 2.7.1.37)
 DE (MAPK/ERK kinase kinase 6) (MEK kinase kinase 6) (MEKKK 6)
 DE (Mishapen/NIK-related kinase) (GCK family kinase MINK).
 GN MAP4K6 OR MINK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), FUNCTION, TISSUE SPECIFICITY,
 RP AND DEVELOPMENTAL STAGE.
 RC STRAIN=ICR; TISSUE=Brain;
 RX MEDLINE=20175403; PubMed=10708748;
 RA Dan I., Watanabe N.M., Kobayashi T., Yamashita-Suzuki K., Fukagaya Y.,
 RA Kajikawa E., Kimura W.K., Nakashima T.M., Matsumoto K.,
 RA Ninomiya-Tsuji J., Kusumi A.;
 FT "Molecular cloning of MINK, a novel member of mammalian GCK family
 FT kinases, which is up-regulated during postnatal mouse cerebral
 FT development.";
 FT FEBS Lett. 469:19-23 (2000).
 [2]
 RP SEQUENCE OF 148-1308 FROM N.A. (ISOFORM 3).
 RC TISSUE=Brain;
 RX MEDLINE=23388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Hellon E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Rodriguez R.C., Touchman J.W., Green E.D., Dickson M.C.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- FUNCTION: Serine/threonine kinase that may play a role in the
 CC response to environmental stress. Appears to act upstream of the
 CC c-Jun N-terminal pathway.
 CC -!- FUNCTION: May play a role in the development of the brain.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphorylated protein.
 CC -!- COFACTOR: Magnesium.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Name=2; Synonyms=MINK-2;
 CC IsoId=Q9JW52-1; Sequence=Displayed;

CC Name=1; Synonyms=MINK-1;
 CC IsoId=Q9JW52-2; Sequence=VSP_007062;
 CC Name=3;
 CC IsoId=Q9JW52-3; Sequence=VSP_007061;
 CC Note=No experimental confirmation available;
 CC TISSUE SPECIFICITY: Appears to be ubiquitous, expressed in all
 CC tissue types examined. Highly expressed in the brain, moderately
 CC expressed in kidney and spleen, low levels present in heart and
 CC skeletal muscle. Isoform 2 is more abundant in the brain than
 CC isoform 1.
 CC -!- DEVELOPMENTAL STAGE: Up-regulated during post-natal brain
 CC development.
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
 CC STE20 subfamily.
 CC -!- SIMILARITY: Contains 1 CNH domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB041925; BA94837.1; -;
 CC EMBL; AB035697; BA90752.1; -;
 CC EMBL; BC011346; AAH11346.1; -;
 CC MGD; MGI:1355329; Map4K6.
 CC GO; GO:0005524; F:ATP binding; IDA.
 CC GO; GO:0004674; P:protein serine/threonine kinase activity; IDA.
 CC GO; GO:0007275; P:development; ISP.
 CC GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
 CC GO; GO:0007243; P:protein kinase cascade; IDA.
 CC GO; GO:0006950; P:response to stress; IDA.
 CC InterPro; IPR001180; Citron.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR008271; Ser_thr_kin_AS.
 CC InterPro; IPR002290; Ser_thr_kinase.
 CC InterPro; IPR001245; Tyr_kinase.
 CC Pfam; PF00780; CNH; 1.
 CC Pfam; PF00069; pkinase; 1.
 CC ProDom; PD000001; Prot_kinase; 1.
 CC SMART; SM00036; CNH; 1.
 CC SMART; SM00220; S_TKc; 1.
 CC SMART; SM00219; TyrKc; 1.
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 CC Transferase; Serine/threonine-protein kinase; ATP-binding;
 KW Alternative splicing.
 KW DOMAIN 25 289 PROTEIN KINASE.
 FT DOMAIN 995 1286 CNH.
 FT NP_BIND 31 39 ATP (BY SIMILARITY).
 FT BINDING 54 54 ATP (BY SIMILARITY).
 FT ACT_SITE 153 153 BY SIMILARITY.
 FT VARSPLIC 698 698 A -> APRPSNSAQVILQRAERGTGPKPGPPAPGPPN
 FT AS (in isoform 3).
 FT /FTid=VSP_007061.
 FT Missing (in isoform 1).
 FT /FTid=VSP_007062.
 FT I -> V (IN REF. 2).
 FT F -> L (IN REF. 2).
 FT E -> R (IN REF. 2).
 FT RSQAG -> AKPSS (IN REF. 2).
 FT MISSING (IN REF. 2).
 FT S -> T (IN REF. 1).
 FT CONFLICT 840 840
 FT SEQUENCE 1308 AA; 147323 MW; 0F366A4E0A87C328 CRC64;
 SQ
 Query Match 93.3%; Score 6464; DB 1; Length 1308;
 Best Local Similarity 92.3%; Pred. No. 7.7e-189;
 Matches 1241; Conservative 17; Mismatches 17; Indels 70; Gaps 6;
 QY 1 MGDPAFARSLDDIDLSALRDPAGIFELVEVVGNTGYGVYGRHVKTGQLAAIKVMDVTE 60

Db 721 PLQRTSSGSSSSSTPSSQSGSQSGSSTRTVRANSKSGSVLPHPBA-KV 779
QY 768 KPDHHR--SRGPRADPVL-----LKERTLDEAPPPKAMDYSSSSSERVSS 813
Db 780 KPEESDITRPRPSYKKAIDDLTALAKEURELRIETNRPKVKVDTSSSESS 839
QY 814 EDDEEGEGGPAAGS---RDTP-----GGRSDGDDTDSVTMVVHDVEEITGTQPPYGG-- 863
Db 840 EEEEDGESETHDGTVAVSDIPRLIPTGAPGSNEQVNVGMVGTGHE--TSHADSPSGSI 897
QY 864 ----GTWVQRTPEERENLHASNGVT----NLPDVVQPSHS-----PTENKSGQSPSKD- 912
Db 898 SREGTLMIRESGKRSKSHSDNGFAGHINLPDLVQOQSHSPAGTPTTEGLGRVSTHSQEM 957
QY 913 GSGDYQSRGLVAPKSGSFTMFVDLGIYOPGGS----GDSIPITALVGEGTRLDQLOQY 967
Db 958 DSGTEYGMG---SSYKASTTTPFVPRVYQTSPTDEDEDESSAAALFTSELLRQOAKL 1014
QY 968 -DVRKGSVVNVNPTNRAHSETPEIRKYKRFENSEILCAALGVNLLVGTENGLMLLDRS 1026
Db 1015 NEARKISVVNVNPTNIRPHSDTPEIRKYKRFENSEILCAALGVNLLVGTENGLMLLDRS 1074
QY 1027 GQKVVGLIGRRRFOQMDVLEGLNLLITISGRNKLRYVYLSWLRNKLHNDPEYKQOG 1086
Db 1075 GQKVVNLRNRRFOQMDVLEGLNVLVTISGKNKLRYVYLSWLRNKLHNDPEYKQOG 1134
QY 1087 WTTVGDMECGHVRVVKYERIKFLVIALKSSVEVYAWAPKPYHKFMFKSPADLPHRPILL 1146
Db 1135 WITVGDLECIYKVKYERIKFLVIALKNAVEIYAWAPKPYHKFMFKSPADLQHPILL 1194
QY 1147 VDLTVBEGORLKVYIGSSAGFHAVDVDSGNSDIYIPVHQIQTIPHAIIFLPNTDGMEM 1206
Db 1195 VDLTVBEGORLKVIFGSHGTGFHVDVDSGNSDIYIPSHIQNITPHAIVILPKTDGMEM 1254
QY 1207 LLYCDEBEGVNTYGRRIKDVVLQWGEEMPTSVAYTCSNQIMGWGKAIEIRSVETGHLDG 1266
Db 1255 LVCYDEBEGVNTYGRITKDVVLQWGEEMPTSVAYTHSNQIMGWGKAIEIRSVETGHLDG 1314
QY 1267 VFMHKAQRLKFLCERNDKVFFASVRSGGSSQVYFMTLRNRCIMMW 1312
Db 1315 VFMHKAQRLKFLCERNDKVFFASVRSGGSSQVYFMTLRNRCIMMW 1360

RESULT 4
M4K4 HUMAN
ID M4K4 HUMAN STANDARD; PRT; 1239 AA.
AC Q95819; O75172; Q9NS77;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 4 (EC 2.7.1.37)
DE (MAPK/ERK kinase kinase 4) (MEK kinase kinase 4) (MEKKK 4)
DE (HPK/GCK-like kinase HGK) (Nck interacting kinase).
GN NAKP44 OR HGK OR NIK OR KIAA0687.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 4 AND 5), FUNCTION, AND TISSUE SPECIFICITY.
RC TISSUE=Macrophage;
RX MEDLINE=99107863; PubMed=9890973;
RA Yao Z., Zhou G., Wang X.S., Brown A., Diener K., Gan H., Tan T.-H.;
RT 'A novel human STE20-related protein kinase, HGK, that specifically activates the c-Jun N-terminal kinase signaling pathway.';
RL J. Biol. Chem. 274:2118-2125 (1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Glialblastoma;
RX MEDLINE=22499944; PubMed=12612079;
RA Wright J.H., Wang X., Manning G., LaMere B.J., Le P., Zhu S.,
RA Khatri D., Flanagan P.M., Buckley S.D., Whyte D.B., Howlett A.R.,

RA Bischoff J.R., Lipson K.E., Jallal B.;
RT "The STE20 kinase HGK is broadly expressed in human tumor cells and can modulate cellular transformation, invasion, and adhesion.";
RL Mol. Cell. Biol. 23:2069-2082 (2003).
RN [3]
RP SEQUENCE OF 38-1239 FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=98403880; PubMed=9734811;
RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
RA Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. X. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";
RL DNA Res. 5:169-176 (1998).
RN [4]
RP SEQUENCE OF 262-1239 FROM N.A. (ISOFORM 3).
RC TISSUE=Testis;
RX Duesterhoeft A., Lauber J., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 700-1239 FROM N.A.
RC TISSUE=Brain;
RA Saito T., Seki N., Hori T.;
RT "Isolation, expression profile and chromosome assignment of a novel serine/threonine kinase gene.";
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [6]
RP INTERACTION WITH SPG3A.
RX MEDLINE=22275851; PubMed=12387898;
RA Luan Z., Zhang Y., Liu A., Man Y., Cheng L., Hu G.;
RT "A novel GTP-binding protein hGBP3 interacts with NIK/HGK.";
RL FEBS Lett. 530:233-238 (2002).
CC -|- FUNCTION: Serine/threonine kinase that may play a role in the response to environmental stress and cytokines such as TNF-alpha. Appears to act upstream of the c-jun N-terminal pathway.
CC -|- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -|- COFACTOR: Magnesium.
CC -|- SUBUNIT: Interacts with the SH3 domain of the adapter proteins Nck (By similarity). Binds, via its CNH regulatory domain, to the N-terminal region of SPG3A.
CC -|- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=5;
CC Name=1; Synonyms=Tumor-associated;
CC IsoId=O95819-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O95819-2; Sequence=VSP_007054, VSP_007057;
CC Name=3;
CC IsoId=O95819-3; Sequence=VSP_007056, VSP_007057;
CC Name=4; Synonyms=HGK-S;
CC IsoId=O95819-4; Sequence=VSP_007055, VSP_007057,
CC Name=5; Synonyms=HGK-L;
CC IsoId=O95819-5; Sequence=VSP_007054, VSP_007055, VSP_007056, VSP_007057, VSP_007058;
CC -|- TISSUE SPECIFICITY: Appears to be ubiquitous, expressed in all tissue types examined. Isoform 5 appears to be more abundant in the brain, isoform 4 is predominant in the liver, skeletal muscle and placenta.
CC -|- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. STE20 subfamily.
CC -|- SIMILARITY: Contains 1 CNH domain.
CC -|- CAUTION: Ref. 4 sequence differs from that shown due to a frameshift in position 1151.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF096300; AAD16137.1; -.


```
Db 702 KSEGP--SPRGSAKPPDDKKVFRSLKPGAGEVDLTALAKELRAVEDVRPPHKVTDYS 759
Qy 805 SSSEV-----ESSEDEEGEGCGPAGSRDTPGCR-----SGDSDTSVSTVWVHDVEEITG 856
Db 760 SSSESGTTDEEEDVEQEGADDSGPDTRAASSPNLSNGETESVKTMTIVHDDVSESP 819
Qy 857 TOPPYGGGTVMVORTPEERNLLHADSNGYTNLPDVQPSHPSTENSKQSPSPKDGSGD 916
Db 820 AMTPSKEGTLVROT-----QSASST----- 840
Qy 917 YQSRGLVAPGKSSFTMFVDLGIYQGGGDSIPITALVG--GEGTRLDQLQYD--VRKGS 973
Db 841 -----LQHKSSSTFTPIDRLLQISPSGT-TVTSVVGSCDGLRPEAIRQDPTKGS 894
Qy 974 VVNVPTNTRAHSETPRIKRYKRNSEBILCAALGVNLLVGTENGLMLLDRSGGQVYG 1033
Db 895 VVNVPTNTRPQSDTPRIKRYKRNSEBILCAALGVNLLVGTESGLMLLDRSGGQVYP 954
Qy 1034 LIGRRRFOQMDVLEGLNLLITISGRNKLRYVYLSWLNKILHNDPEVEKKQGTWTTGDM 1093
Db 955 LISRRRFOQMDVLEGLNLLVITSGKKDLRYVYLSWLNKILHNDPEVEKKQGTWTTGDL 1014
Qy 1094 ECGHYRVVYKRIKFLVIALKSSVEVYVAPKPYHKFMFKSPADLRPLLDVLTVEE 1153
Db 1015 ECGHYRVVYKRIKFLVIALKSSVEVYVAPKPYHKFMFKSPGELLHKPLLDVLTVEE 1074
Qy 1154 GQRLKVIYGSAGHAFVDSGNSYDIYPIVHIQOITPHAIIFLPNTDGMELLCYBDE 1213
Db 1075 GQRLKVIYGSAGHAFVDSGNSYDIYPIVHIQOITPHAIIFLPNTDGMELLCYBDE 1134
Qy 1214 GYVNTYGRITKDVVLQNGEMPTSVAYICSNQIMGEGKAIEIRSVETGHLGDVFMHKA 1273
Db 1135 GYVNTYGRITKDVVLQNGEMPTSVAYIRSNQTMGEGKAIEIRSVETGHLGDVFMHKA 1194
Qy 1274 QRLKFLCERNKVFASVRSGSSGVYFMTLNRCIMW 1312
Db 1195 QRLKFLCERNKVFASVRSGSSGVYFMTLGRSTLSW 1233
```

RESULT 6

```
TN1K_MOUSE
ID TN1K_MOUSE STANDARD; PRT; 916 AA.
AC P83510;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Trf2 and NCK interacting kinase (EC 2.7.1.37) (Fragments).
GN TN1K OR KIAA0551.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=C57BL/6J, and NOD; TISSUE=Hypothalamus, and Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Nikaide I., Osato N., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Sult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Guscinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Peste G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shlmada K.,
```

```
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Haseizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [2]
RP SEQUENCE OF 359-842 FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=22579291; PubMed=12693553;
RA Okazaki M., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:35-48(2003).
CC -1- FUNCTION: Stress-activated serine/threonine kinase that may play a
CC role in the response to environmental stress. Appears to act
CC upstream of the c-jun N-terminal pathway. May play a role in
CC cytoskeletal regulation (by similarity).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COPACTOR: Magnesium (By similarity).
CC -1- SUBUNIT: Interacts with TRAF2 and NCK (By similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P83510-1; Sequence=Displayed;
CC Note=No experimental confirmation available;
CC Name=2;
CC IsoId=P83510-2; Sequence=VSP_007351;
CC Note=No experimental confirmation available;
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC STE20 subfamily.
CC -1- SIMILARITY: Contains 1 CNH domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AK039113; BAC30241.1; -
CC EMBL; AK041777; BAC31061.1; -
CC EMBL; AK088459; BAC40365.1; -
CC EMBL; AK122306; BAC65588.1; -
CC MGD; MGI:2444917; C53000801Srik.
CC GO; GO:0004674; F:protein serine/threonine kinase activity; ISS.
CC GO; GO:0007243; P:protein kinase cascade; ISS.
CC GO; GO:0006445; P:regulation of translation; ISS.
CC GO; GO:0006950; P:response to stress; ISS.
CC InterPro; IPR001180; Citron.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_thr_pkin_AS.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00780; CNH; 1.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00036; CNH; 1.
CC SMART; SM00220; S_TKc; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00011; PROTEIN KINASE DOM; 1.
CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC ATP-binding; Serine/threonine-protein kinase; Transferase;
KW
```

KW	Alternative splicing.	
FT	DOMAIN	25 289
FT	NP_BIND	31 39
FT	BINDING	54 54
FT	ACT_SITE	153 153
FT	NON_CONS	327 328
FT	DOMAIN	603 894
FT	VARSPLIC	519 555
FT	SEQUENCE	916 AA; 103344 MW; 7C40997942B5347C CRC64;
FT	Query Match	47.4%; Score 3286; DB 1; Length 916;
FT	Best Local Similarity	51.2%; Pred. No. 4.4e-93;
FT	Matches	689; Conservative 76; Mismatches 117; Indels 464; Gaps 16;
QY	1	MCDPAPASLDDIDLSALRDPAGIFELVELVVGNGTYGQYKGRHVKTQGLAAIKMDYTE 60
DB	1	MASDSPARSLEIDLSALRDPAGIFELVELVVGNGTYGQYKGRHVKTQGLAAIKMDVTG 60
QY	61	DEEEIKQEIINMKYSHRNATATYGAFIKKSPGNDQDLWLVMFECAGSVTLVKNT 120
DB	61	DEEEIKQEIINMKYSHRNATATYGAFIKKSPGNDQDLWLVMFECAGSVTLVKNT 120
QY	121	KGNALEKDCIAYICREILRGLAHLAHKVIHRDIKQNVLLTENAELVDFGVSQAQLDR 180
DB	121	KGNTLKEEWIAYICREILRGLSHLHQHKVHRDIKQNVLLTENAELVDFGVSQAQLDR 180
QY	181	TVGRNTFTGTWYMAPEVIACDENPDATYDRSDIWSLIGITAIEMAGAPPLCDMPMR 240
DB	181	TVGRNTFTGTWYMAPEVIACDENPDATYDFKSLWSLIGITAIEMAGAPPLCDMPMR 240
QY	241	ALFLIPRNPPLKSKKSKKFIIDTCLIKTYLSRPTEQLLKFPIRQDPTEROVRI 300
DB	241	ALFLIPRNPAPLKKSKKSKKQSFIESCLVKNHSQRPAEQMLKHPFIRQDPTEROVRI 300
QY	301	QLKDHIDRSKRKGKETEYSGSEEDSHGEGEPSSIMNVPGESTLRLREFRLQQ 360
DB	301	QLKDHIDRTKKRGEKDETEYSGSECK-----SEGSP----- 334
QY	361	ENKNSSEALKQOQQLQOQOQDPEAHIKHLLHORRRIEQQEERRRVEEQRRERQK 420
DB	335	----- 334
QY	421	LOEKEQRRLEDQALREERERRQAEQYKQLEBQQRQSERLQRLQOEHAFLKSLQ 480
DB	335	----- 334
QY	481	00000000LQK00000LPGDRKPLHYHGRGNPADKPAREVEERTRMNKQNSPLAKS 540
DB	335	-----VLP----- 337
QY	541	KPGSTGPBPPIPOASGPPGGLSQTTPMQRPVEPQEGPHKSLQDQPTNLAAFPASHDPD 600
DB	338	-----HEP- 340
QY	601	PAIPATATPSARGAVIRQNSDPTSEGPCSPGNPAWVRPDNEAPKVPQRTSSIALTN 660
DB	341	----- 340
QY	661	TSGAGSRPAQAVRARPRNSAWQIYLQRAERGTTPKPPGPPAQPPGPPNASSNPDLRRS 720
DB	341	----- 340
QY	721	DPGERSDVLPAASHGLPQAGSLERNVRGASSKLDSPVLSPGNKAKYDPDHR--SRPGR 778
DB	341	-----SKVKEPSRDIITPSR 356
QY	779	PADFVL-----LKERTLDAPRPKKAMDYSSSSEVESSEDEBEEGGPAEGS---RDT 831
DB	357	PADLTALAKELRELRIETNRLPKKVTYSSSSESESESESESESETHDGTAVVSDI 416
QY	832	P-----GGRSDGDTDSVSTMVVDHVEITGTQPPYGG-----GTMVQRTPEERNLLHA 881

Db	417	PRLLPTGAPGNEQYNNMGVGTGHLG--TSHADTFGGSISREGTLMIRETAEKRSCHS 474
QY	882	DSNGYT---NLDPVVPQSHS---PTENSKQSQPSK---DSGQDYQSRGLVKAPGKSFT 932
Db	475	DSNGFAGHINLPDLVQOQSHSPAGTFTGLGRVSTHSGEMDSGABYG---IGSSTKASFT 530
QY	933	MFVDLGIYQPGGS-----GDSIPITALVGGSGTGLDQLOY--DVRKGSVNVNPNPTNTRAHS 986
Db	531	PFVDPRVYQTSPTDEDEDESSAAALFTSELLAQEQAKLNEARKISVVNVNPNTRPHS 590
QY	987	ETPEIRKYKRFNSEILCAALWGNLLVGTENGMLLDRSQGKYGLIGRRRFOQMDVL 1046
Db	591	DTPEIRKYKRFNSEILCAALWGNLLVGTENGMLLDRSQGKYGLIGRRRFOQMDVL 650
QY	1047	EGLNLLTTSKRNKRLAVVYLSWLNKILHNDPVEKKQGTWTTVDMEGGCHYVVKYER 1106
Db	651	EGLNVLTISGKNKRLVYVYLSWLNRLHNDPVEKKQGTWTTVDMEGGCHYVVKYER 710
QY	1107	IKFLVIALK3SVYAWAPKPYHKFMFKSFADLPHPRLVLDLTVEGQRLKVIYGSAG 1166
Db	711	IKFLVIALKNAVEIYANAPKPYHKFMFKSFADLPHPRLVLDLTVEGQRLKVIYGSHTG 770
QY	1167	PHAVDVSGNSYDIYIPVHIQSQITPHAIIFLPNTDGMELLCYEDGEGVYVNTYGRITKD 1226
Db	771	PHVIDVSGNSYDIYIPSHIQGNITPHAVILPKTDGMEMLVCYEDGEGVYVNTYGRITKD 830
QY	1227	VVLQMGEMPTSVAVICSNQIMGGEKAIEIRSVETGHLGDFVFMHKAQRLKFLCERNDKV 1286
Db	831	VVLQMGEMPTSVAVIHSNQIMGGEKAIEIRSVETGHLGDFVFMHKAQRLKFLCERNDKV 890
QY	1287	FFASVRSGSSQVYFMTLNRNCIMNW 1312
Db	891	FFASVRSGSSQVFFMTLNRNSMNW 916

RESULT 7

M15 CAEL

ID M15 CAEL STANDARD; PRT; 1080 AA.

AC Q23356;

DT 15-DEC-1998 (Rel. 37, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Serine/threonine-protein kinase mig-15 (EC 2.7.1.-).

GN MIG-15 OR ZC504.4.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RA	Kershaw J.;	
RL	Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.	
RN	(2)	
RP	REVISIONS.	
RA	Durbin R.;	
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.	
RN	(3)	
RP	IDENTIFICATION.	
RA	Zhu X., Hedgecock E.;	
RT	"mig-15 encodes a novel Ser/Thr protein kinase of the Ste-20/p65PAK family.";	
RL	(In) Worm Breeder's Gazette 14(5):76(1997).	
CC	-!- FUNCTION: IMPORTANT IN SEVERAL DEVELOPMENTAL PROCESSES INCLUDING EPIDERMAL DEVELOPMENT, Q NEUROBLAST MIGRATIONS AND MUSCLE ARM TARGETING.	
CC	-!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. STE20 subfamily.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its	

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; Z50029; CAA90344.1; -
DR PIR; T27622; T27622.
DR WormPep; ZC504.4; CE19311.
DR InterPro; IPR001180; Citron.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser Thr pkin AS.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 21 281 PROTEIN_KINASE.
FT NP_BIND 27 35 ATP (BY SIMILARITY).
FT BINDING 50 50 ATP (BY SIMILARITY).
FT ACT_SITE 151 151 BY SIMILARITY.
FT DOMAIN 757 1060 CNH.
SQ SEQUENCE 1080 AA; 120785 MW; 4DAB06070FE26F63 CRC64;

Query Match 38.2%; Score 2646; DB 1; Length 1080;
Best Local Similarity 43.3%; Pred. No. 1e-73;
Matches 590; Conservative 156; Mismatches 268; Indels 350; Gaps 35;

QY 10 LDDIDLSALRDPAGIFELVVGNGTYGVYKGRHVKTGQLAAIKWMDVTEDEEBEIKOE 69
DB 6 LDEIDLNSLRDPAGIFELVVGNGTYGVYKGRHVKTGQLAAIKWMDVTEDEEBEIKOE 65

QY 70 INMLKKYSHHRIATYGAFFKSP--FGNDQLWLWVEFCAGSVTDLVKNTKGNALKE 127
DB 66 INMLKKSHHRIATYGAFFKSPSTGKHDLWLWVEFCAGSVTDLVKNTKGSLKE 125

QY 128 DCIAYICREILRLGLAHKVIHRIKQNVLLTENAELVDFGSAQLDRVTGRRNT 187
DB 126 EWAIYICREILRLGLHQSQVIHRIKQNVLLTDSAELVDFGSAQLDRVTGRRNT 185

QY 188 FIGTPYMAPEVIACDENPDATYRSDIWSLGIATAIEMAGAPLCDHMPRALFLIPR 247
DB 186 FIGTPYMAPEVIACDESPEATYDSK-----ITALEMAGHPPLCDHMPRALFLIPR 238

QY 248 NPPRLK-SKKWKKFIDFIDTCLIKTYLSRPTQLKFPFIRDOPTQVIRQLKDIH 306
DB 239 NPPPKLKKKKWTKKFETFIETLVKYHORPYTGALLRHPPFIKEQPEQTIRHSKEH 298

QY 307 DRSRKRGEKETETEYSGSEEDSDSHGESEPSIMN-----VPGESTLRRREFLRQ 359
DB 299 DNRN--RVKQDADYEYSGSEDESPNRRPSSGIRDDSESSMIPMDNLRKGFQKLQ 356

QY 360 QENKSNSEALKQ--QOOLQOQQOQDPEAHIKHLHQRRRIEQQE-----ERRRVE 409
DB 357 ESSRGAEPGAQQLRLRQPPAP-----FQYQSRVYVPRRESSVKKLRAVSSRCAA 410

QY 410 EQQRERQRKLQEKQQRRLD-----MQALRREERQAEQEQQYKKQLEEQRQS 462
DB 411 DGPRHSPASRPRPVSHHRSQQSHAPAPHLADLANYEKKRRRSERE----- 456

QY 463 ERLQRLQOQEHAYLSLQOQQOQQOQLQOQQOQLPGDRKPLHYHGRGNWPNADKPWARE 522
DB 457 ERRERERQAHAM----- 469

QY 523 VEERTRMNKQNSPLAKSPGSGTGPPIPPQASPGPGPLSGTTPMWRP-----VEPQEG 577
DB 470 -----PIARVASVPAP-QQSRKMSSEPLITHVKPED- 500

QY 578 PHKSLODQPTRNLAAFPASHDPDAIPAPTATPSARGAVIRQNSDPTSEGPGSPNPPAW 637

Db 501 -----LDVLASELSKMGCHH-----NORSREE-----SMSPP-- 527
QY 638 VRPDNEAPPKVPQRTSTIATLANTSGAGSRPAQAVRARPRNSAWQIY--LORRAERG 695
Db 528 -----PPAPPPREASISSTIDIVG-----ELDNGADAEMDLDIMWNGE-GT 571
QY 696 PKPPGPPAQPQPPNASSN--PDLRRSDPGWERSDVLPAHSHGLPOAGSLERNVRV 753
Db 572 LRGNKPL-PPTPTDGTENTLVSDVRNNG-----NSGH-----GAYKGKKI-- 612
QY 754 KLDSSPVLSPGNKAKPDDHRSRPGPADFVLLKERTLDEAPRPPKAMDYSSSEEVES 813
Db 613 -----PEIRGIIISLDD-----DS 627
QY 814 EDDEEGEGGPAEGSRDTPGGRSDGTSVSMVVDHVEITGTTPPYGGGTVMVQRTPE 873
Db 628 DSDNEEG-----NEPL 638
QY 874 EERNLLHADNGYTNLPDVVQPSHSPTENSKGQSPKSGDYGQSRGLVAKPKSSE-- 931
Db 639 MFKPINASSRG--ALPDLFPKLPQLRRQINDQTRQMSDDRADPQPNFQNSDRSIQH 696
QY 932 -----TMFVDLGIYQPGSGSDSIPTALVGEGT-----RLDQLQYDVRKGSV 975
Db 697 SFSNRDEKSFVG--YFGGAG-----AGGGTVNRPGRPDINQVQ-----V 736
QY 976 NVNPTN--TRAHSETPEIRYKRRFNSSEILCAALWGVNLLVGTENGLMLDRSQGKYV 1033
Db 737 NVTNSNGTPAENDAPAIRYKFKFSGEILCAALWGVNLLIGTDSGLMLDRSQGKYV 796
QY 1034 LIGRRRQOMDVLEGLNLLITISGRKRLRVYLSWLRNKLHND-----PEVEKKQGW 1088
Db 797 LISRRRQOMDVLEGLNLLITISGRKRLRVYLSWLRNKLHND-----PEVEKKQGW 856
QY 1089 TVGDMCCGHRVYVYKRIKFLVIALKSSVEVYAWAPKPKVMAKFSADLPHRPLVD 1148
Db 857 NVGDLQAIHFKVYRIKFLVVGLESEIETAWAPKPKVMAKFSADLPHRPLVD 916
QY 1149 LTVEEGRLKVIYSSAGFAHVDVDSGNSYDIYIPVHIQSQITPHAIIFLPTDGMEMLL 1208
Db 917 LTVEDNARLKVYSGTGGFHAIDLDSAAYDIYTPAQSGQTTTTHFVLPNSNGMQLL 976
QY 1209 CYDEGVYVNTYGRRIKDVVLQWGMPTSVAYICSNQIMGWGEKAIIRSVETGHLQGV 1268
Db 977 CYDNEGVYVNTYGRMTKXVVLQWGMPTSVAYISGTQIMGWGNKAIIRSVETGHLQGV 1036
QY 1269 MHKBAORLKLCEKNDKVPFASVRSVGSSGVYFMTLRNCIMMW 1312
Db 1037 MHKKAQKLKLCERNDKVPFASVRSVGSSGVYFMTLRNCIMMW 1080

RESULT 8
M4K3 HUMAN
ID M4K3_HUMAN STANDARD; PRT; 894 AA.
AC Q8IVH8; Q8IVH7; Q9UDMS; Q9Y6R5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 3 (EC 2.7.1.37)
DE (MAPK/ERK kinase kinase kinase 3) (MEK kinase kinase 3) (MEKKK 3)
DE (Germinal center kinase related protein kinase) (GLK).
GN MAP4K3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2), FUNCTION, TISSUE SPECIFICITY, AND
RP MUTAGENESIS OF LYS-48.
RC TISSUE=Macrophage, and Skeletal muscle;
RX MEDLINE=97420743; PubMed=9275185;
RA Diener K., Wang X.S., Chen C., Meyer C.F., Keesler G., Zukowski M.,

RA Tan T.-H., Yao Z.;
RT "Activation of the c-Jun N-terminal kinase pathway by a novel protein
RT kinase related to human germlinal central kinase.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:9687-9692(1997).
RN (2)
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
RA Gorry M.C., Zhang Y., Marks J.J., Suppe B., Hart S., Cortelli J.,
RA Pallos D., Hart T.C.;
RT "Physical/genetic map of the 2p22-2p21 region on chromosome 2.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN (3)
RP SEQUENCE OF 1-712 FROM N.A. (ISOFORM 1).
RA Edwards J., Wohldmann P., Hawkins M., Harkins R.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May play a role in the response to environmental stress.
CC Appears to act upstream of the c-jun N-terminal
CC pathway.
CC
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a
CC phosphoprotein.
CC
CC -!- COFACTOR: Magnesium.
CC
CC -!- SUBUNIT: Interacts with SH3GL2. Interaction appears to regulate
CC MAPK3-mediated JNK activation (By similarity).
CC
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q8IVH8-1; Sequences=Displayed;
CC Note=No experimental confirmation available;
CC Name=2;
CC IsoId=Q8IVH8-2; Sequences=VSP_007052;
CC Name=3;
CC IsoId=Q8IVH8-3; Sequences=VSP_007053;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed in all tissues
CC examined, with high levels in heart, brain, placenta, skeletal
CC muscle, kidney and pancreas and lower levels in lung and
CC liver.
CC
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC STP20 subfamily.
CC
CC -!- SIMILARITY: Contains 1 CNH domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AF000145; AAC15472.1; -.
CC EMBL; AF445413; AAN75849.1; -.
CC EMBL; AF445385; AAN75849.1; JOINED.
CC EMBL; AF445386; AAN75849.1; JOINED.
CC EMBL; AF445387; AAN75849.1; JOINED.
CC EMBL; AF445388; AAN75849.1; JOINED.
CC EMBL; AF445390; AAN75849.1; JOINED.
CC EMBL; AF445391; AAN75849.1; JOINED.
CC EMBL; AF445392; AAN75849.1; JOINED.
CC EMBL; AF445393; AAN75849.1; JOINED.
CC EMBL; AF445394; AAN75849.1; JOINED.
CC EMBL; AF445395; AAN75849.1; JOINED.
CC EMBL; AF445396; AAN75849.1; JOINED.
CC EMBL; AF445397; AAN75849.1; JOINED.
CC EMBL; AF445398; AAN75849.1; JOINED.
CC EMBL; AF445399; AAN75849.1; JOINED.
CC EMBL; AF445400; AAN75849.1; JOINED.
CC EMBL; AF445401; AAN75849.1; JOINED.
CC EMBL; AF445402; AAN75849.1; JOINED.
CC EMBL; AF445403; AAN75849.1; JOINED.
CC EMBL; AF445404; AAN75849.1; JOINED.
CC EMBL; AF445405; AAN75849.1; JOINED.
CC EMBL; AF445406; AAN75849.1; JOINED.
CC EMBL; AF445407; AAN75849.1; JOINED.
CC EMBL; AF445408; AAN75849.1; JOINED.
CC EMBL; AF445409; AAN75849.1; JOINED.
CC EMBL; AF445410; AAN75849.1; JOINED.
CC EMBL; AF445411; AAN75849.1; JOINED.
CC EMBL; AF445412; AAN75849.1; JOINED.
CC EMBL; AC007684; AAF19240.1; -.
CC HSP; P24941.1B38.
CC Genew: HGNC:6865; MAP4K3.
CC MIM: 604921; -.
CC GO; GO:000524; F:ATP binding; IDA.
CC GO; GO:0004674; F:protein serine/threonine kinase activity; IDA.
CC GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
CC GO; GO:0007243; P:protein kinase cascade; IDA.
CC GO; GO:0006950; P:response to stress; IDA.
CC InterPro; IPR001180; Citron.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_thr_pkin_AS.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00780; CNH; 1.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00036; CNH; 1.
CC SMART; SM00220; S_TKc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE NEG.
CC ATP-binding; Transferase; Serine/threonine-protein kinase;
CC Alternative splicing.
CC DOMAIN 16 273 PROTEIN KINASE.
CC FT NP_BIND 22 30 CNH.
CC FT BINDING 48 48 ATP.
CC FT ACT_SITE 136 136 BY SIMILARITY.
CC FT VARSPPLIC 1 12 MNPFGDLRRNP -> MA (in isoform 2).
CC FT VARSPPLIC 352 372 Missing (in isoform 3).
CC FT FTID=VSP_007053.
CC FT FTID=VSP_007053.
CC FT MUTAGEN 48 48 K->E: LOSS OF KINASE ACTIVITY AND ABILITY
CC FT CONFLICT 392 392 N -> D (IN REF. 1; AAC15472).
CC FT CONFLICT 392 392 N -> D (IN REF. 1; AAC15472).
SQ SEQUENCE 894 AA; 101315 MW; 6EB77BBB34E5B733 CRC64;
Query Match 10.7%; Score 738; DB 1; Length 894;
Best Local Similarity 29.1%; Pred. No. 3.2e-16;
Matches 200; Conservative 97; Mismatches 207; Indels 184; Gaps 22;
Qy 14 DLSALRDPAGIFELVGVNGTYGVYKGRHRTGTGQLAAIKVMDVTEDEEEIKQEIINML 73

DR EMBL; AF445409; AAN75849.1; JOINED.
DR EMBL; AF445410; AAN75849.1; JOINED.
DR EMBL; AF445411; AAN75849.1; JOINED.
DR EMBL; AF445412; AAN75849.1; JOINED.
DR EMBL; AF445413; AAN75850.1; -.
DR EMBL; AF445385; AAN75850.1; JOINED.
DR EMBL; AF445386; AAN75850.1; JOINED.
DR EMBL; AF445387; AAN75850.1; JOINED.
DR EMBL; AF445388; AAN75850.1; JOINED.
DR EMBL; AF445390; AAN75850.1; JOINED.
DR EMBL; AF445391; AAN75850.1; JOINED.
DR EMBL; AF445392; AAN75850.1; JOINED.
DR EMBL; AF445393; AAN75850.1; JOINED.
DR EMBL; AF445394; AAN75850.1; JOINED.
DR EMBL; AF445395; AAN75850.1; JOINED.
DR EMBL; AF445397; AAN75850.1; JOINED.
DR EMBL; AF445398; AAN75850.1; JOINED.
DR EMBL; AF445399; AAN75850.1; JOINED.
DR EMBL; AF445400; AAN75850.1; JOINED.
DR EMBL; AF445401; AAN75850.1; JOINED.
DR EMBL; AF445402; AAN75850.1; JOINED.
DR EMBL; AF445403; AAN75850.1; JOINED.
DR EMBL; AF445404; AAN75850.1; JOINED.
DR EMBL; AF445405; AAN75850.1; JOINED.
DR EMBL; AF445406; AAN75850.1; JOINED.
DR EMBL; AF445407; AAN75850.1; JOINED.
DR EMBL; AF445408; AAN75850.1; JOINED.
DR EMBL; AF445409; AAN75850.1; JOINED.
DR EMBL; AF445410; AAN75850.1; JOINED.
DR EMBL; AF445411; AAN75850.1; JOINED.
DR EMBL; AF445412; AAN75850.1; JOINED.
DR EMBL; AC007684; AAF19240.1; -.
DR HSP; P24941.1B38.
DR Genew: HGNC:6865; MAP4K3.
DR MIM: 604921; -.
DR GO; GO:000524; F:ATP binding; IDA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IDA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
DR GO; GO:0007243; P:protein kinase cascade; IDA.
DR GO; GO:0006950; P:response to stress; IDA.
DR InterPro; IPR001180; Citron.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE NEG.
KW ATP-binding; Transferase; Serine/threonine-protein kinase;
KW Alternative splicing.
FT DOMAIN 16 273 PROTEIN KINASE.
FT NP_BIND 22 30 CNH.
FT BINDING 48 48 ATP.
FT ACT_SITE 136 136 BY SIMILARITY.
FT VARSPPLIC 1 12 MNPFGDLRRNP -> MA (in isoform 2).
FT VARSPPLIC 352 372 Missing (in isoform 3).
FT FTID=VSP_007053.
FT FTID=VSP_007053.
FT MUTAGEN 48 48 K->E: LOSS OF KINASE ACTIVITY AND ABILITY
FT CONFLICT 392 392 N -> D (IN REF. 1; AAC15472).
SQ SEQUENCE 894 AA; 101315 MW; 6EB77BBB34E5B733 CRC64;


```
Db 6 DLNR-RNQBDFELQIRGSGYGVYKARNVNTGELAAIKVLEPGEDFAVVOOEIIM 64
Qy 74 KYSHRNATYAGAFIKKSPGDDQLWLMVEFCGAGSVTDLVKNTKGNALKEDCIAYI 133
Db 65 MKDCKHPNIVAFGSLRR-----DKLWICMEFCGSGSLQDIYHVT--GPLSELQIAYV 116
Qy 134 CREILRGLAHKVIHRDIKQNVLLTENAEVKLVDFGVSQAQLDRTVGRNTFICTPY 193
Db 117 SRETLQGLYLLHSGKMRDIKGANILLTDNGHVKLADFGVSAQITATIAKRKSPIGTPY 176
Qy 194 WMAPEVIACDENPDATYDYSWISGLTATEMAEAGAPLDMHPRALFLIPRN--PPP 251
Db 177 WMAPEVAVERK--GGYNQLCDLWAVGITALELAELQPPMDLHPMALFLMTKSNFQPP 234
Qy 252 RLKSK-KWSKPFIDFIDTCLIKTYLSRPPTTEQLLKFPIRDQPTQVRQIRLQKHIDRSR 310
Db 235 KLKQKWKSNFPHFVKMALTKNPKRPTAEKLLQHPFVTOHLT-RSLAIELLDKVNPN- 292
Qy 311 KKRGEKETEYSSSEEDSHGEGEPSSIMNVPGESTLRREFLRLQOENKSNSEALK 370
Db 293 -----DHSTYHDFDD-----DPEPLVAVP----- 312
Qy 371 QQQQLQQQQQDPEAHIKHLHQRRRIEEOKEERRRVEEOORREORKEQKEQORRL 430
Db 313 -----HRIHSTSRNVREKTRSEITFGQVKPDPPLRK--ETEPHHEL 352
Qy 431 EDMQA-LRREBE-----RQAREOEYKR-----KQLEORQSERLQRL 469
Db 353 PSDGFLDSEIYYTARNLQLEYGQHGQGYFLGANKSLKSVSELHQHVAHL 412
Qy 470 QOEHAFLSLQOQQQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQOQQO 529
Db 413 EDDEG-----DDDESKHSTLAKIPPLPPPKPSIF-----IPQEMHS 450
Qy 530 NKQNSPLAKSKPGSTGPEPIPOASPGPGPLSOTPMQRPVEPQEGPHKSLQDQPTRN 589
Db 451 TEDENQGTIKKCPMSGSAKP-----SQVPP--RPPPLPPHPKV--ALNGG 494
Qy 590 LAAPFASHPDPPATPAPTATPSARGAVIRQNSD-----PTSEGGPGSPN 633
Db 495 MSSQLNGERD-----GSLCQQQNEHRTGTLNRKEKQDVKPKPSINGLPPTPK 541
Qy 634 -----PPAWVRPD 641
Db 542 VHMGAFCFSKVFNGCPLKIHCASSWINPD 569

RESULT 9
ID M4K3 RAT STANDARD; PRT; 862 AA.
AC Q924I2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Mitogen-activated protein kinase kinase kinase 3 (EC 2.7.1.37)
DE (MAPK/ERK kinase kinase 3) (MEK kinase kinase 3) (MEKK 3)
DE (Germinal center kinase related protein kinase) (GLK) (Fragment).
GN MAP4K3
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI TaxID=10116;
RN [1]_
RP SEQUENCE FROM N.A., AND INTERACTION WITH SH3GL2.
RX MEDLINE=21369947; PubMed=11384986;
RA Ramjaun A.N., Angers A., Legendre-Guillen V., Tong X.-K.,
RA McPherson P.S.;
RT "Endophilin regulates JNK activation through its interaction with the
RT germinal center kinase-like kinase.";
RL J. Biol. Chem. 276:28913-28919(2001);
CC -!- FUNCTION: May play a role in the response to environmental stress.
CC Appears to act upstream of the c-jun N-terminal pathway (By
```

```
similarity).
-!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
-!- COFACTOR: Magnesium (By similarity).
-!- SUBUNIT: Interacts with SH3GL2. Interaction appears to regulate
MAP4K3-mediated JNK activation.
-!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
STE20 subfamily.
-!- SIMILARITY: Contains 1 CNH domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
EMBL; AF312224; AAK53214.1; -.
HSSP; P24941; 1BUH
GO; GO:0005524; F:ATP binding; ISS.
GO; GO:0004674; F:protein serine/threonine kinase activity; ISS.
GO; GO:0006468; F:protein amino acid phosphorylation; ISS.
GO; GO:0007243; P:protein kinase cascade; ISS.
GO; GO:0006950; P:response to stress; ISS.
InterPro; IPR001180; Citron.
InterPro; IPR000719; Prot_kinase.
InterPro; IPR008271; Ser_Thr_pkin_AS.
InterPro; IPR002290; Ser_Thr_pkinase.
Pfam; PF00780; CNH; 1.
Pfam; PF00069; pkinase; 1.
ProDom; PD000001; Prot_kinase; 1.
SMART; SM00036; CNH; 1.
SMART; SM00220; S_TKC; 1.
PROSITE; PS00107; PROTEIN KINASE ATP; 1.
PROSITE; PS00113; PROTEIN KINASE DOM; 1.
PROSITE; PS00108; PROTEIN KINASE ST; FALSE NEG.
KW ATP-binding; Transferase; Serine/threonine-protein kinase.
FT NON_TER 1
FT DOMAIN 5 262 PROTEIN KINASE.
FT DOMAIN 530 842 CNH.
FT NP_BIND 11 19 ATP (BY SIMILARITY).
FT BINDING 34 34 ATP (BY SIMILARITY).
FT ACT_SITE 125 125 BY SIMILARITY.
SQ SEQUENCE 862 AA; 97390 MW; 58013AC3B0A3287F CRC64;

Query Match 10.5%; Score 729; DB 1; Length 862;
Best Local Similarity 30.4%; Pred. No. 5.8e-16;
Matches 200; Conservative 97; Mismatches 203; Indels 158; Gaps 24;

Qy 21 PAGIFELVVGNGTYQGVYKGRHVTKGQLAAIKWMDVTEDEEEIKQEINMLKKYSHHR 80
Db 1 PQEDFELIQIGSGTYGVYKARNVNTGELAAIKVLEPGEDFAVVOQEIMMKCKKHA 60
Qy 81 NIATYGAFIKSPGDDQLWLMVEFCGAGSVTDLVKNTKGNALKEDCIAYICREILRG 140
Db 61 NIVAFGSLRR-----DKLWICMEFCGSGSLQDIYHVT--GPLSELQIAYVRETQQ 112
Qy 141 LAHLHAHKVIHRDIKQNVLLTENAEVKLVDFGVSQAQLDRTVGRNTFICTPYWMAPEVI 200
Db 113 LYLLHSGKMRDIKGANILLTDNGHVKLADFGVSAQITATIAKRKSPIGTPYWMAPEVA 172
Qy 201 ACBENPDATYDYSWISGLTATEMAEAGAPLDMHPRALFLIPRN--PPPRLKSK-K 257
Db 173 AVERK--GGYNQLCDLWAVGITALELAELQPPMDLHPMALFLMTKSNFQPPKLKOKL 230
Qy 258 WSKKPFIDFIDTCLIKTYLSRPPTTEQLLKFPIRDQPTQVRQIRLQKHIDRSRKRGE 317
Db 231 WSNSEFHHFVKMALTKNPKRPTAEKLLQHPFV--TQPLTRSLAIELLDKVNPN----- 281
Qy 318 ETEYEGSGSEEDSHGEGEPSSIMNVPGE--STLRREFLRLQOENKSNSEALKQOQOL 375
Db 282 ----DHSTYHDFDD-----DPEPLVAVPHRIPSTSN-----VREKTRSEINFGQVKF 327
```

```
QY 376 OQQQQRDPBAHKHLL-----HQRRRIEQQEERRRVEEQRRERQKLEKEQ 426
Db 328 DPPLRKETEPH--HELDLQLEYQGHOSNYFLGNGKSLLSVVEB-----LHQGH 376
QY 427 QRRLEDMQALRRREERQERQERQYKQLEERQSERLQROLQOEHAHLKSLQQQQQQ 486
Db 377 VAHLED-----DEGDD-----DSKGSTLKA----- 397
QY 487 QLQKQQQQQLPQDRKPLVHYGKGMNPKAPAWAREVERTRMNQKQNSPLAKSPKSGTG 546
Db 398 -----KVPPPPPP-----KPK-SISIPDTHSSDSNOGTTIKRCPSSGS 435
QY 547 PEPPPIQASGPGPLSPQTPPQRPVQEGPHKSLQDQPTNLAAFPASHDPDPAIPAP 606
Db 436 PAKP-SHVPPRPPPP--RLPP-QKPAVLNG-----VSSFLNGERDGSVHQQ 479
QY 607 TATPSARGAVIQNSD-----PTSEGPGPSN-----PPAWVRPD 641
Db 480 QSEQRGTNLSRKEKDVPKPIGSLNGLPPTPKVHMGCFSKVFNGCPLKIHCAATSWINPD 537

RESULT 10
STK4_HUMAN STANDARD; PRT; 487 AA.
AC Q13043; Q15802; Q9NT24;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine protein kinase 4 (EC 2.7.1.37) (STE20-like kinase
DE MST1) (MST-1) (Mammalian STE20-like protein kinase 1)
DE (serine/threonine protein kinase Krs-2).
GN STK4 OR MST1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95394929; PubMed=7665586;
RA Creasy C.L., Chernoff J.;
RT "Cloning and characterization of a human protein kinase with homology
RT to Ste20.";
RL J. Biol. Chem. 270:21695-21700(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96413604; PubMed=8816758;
RA Taylor L.K., Wang H.C., Erikson R.L.;
RT "Newly identified stress-responsive protein kinases, Krs-1 and Krs-
RT 2.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:10099-10104(1996).
RN [3]
RP SEQUENCE OF 1-435 FROM N.A.
RA Laird G.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- ENZYME REGULATION: THE C-TERMINAL NON-CATALYTIC REGION INHIBITS
CC THE KINASE ACTIVITY.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.
CC -!- PTM: Autophosphorylated on serine and threonine residues.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC STE20 subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U18297; AA083254.1; -.
DR EMBL; U60207; AA017262.1; -.

```

```
DR EMBL; AL109839; CAB89421.1; -.
DR HSSP; P24941; 1HCL.
DR Genew; HGNC:11408; STK4.
DR MIM; 604965; -.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; TAS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser_Thr_pkin_AS.
DR InterPro; IPR002290; Ser_Thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TKC_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 30 281 PROTEIN_KINASE.
FT NP_BIND 36 44 ATP (BY SIMILARITY).
FT BINDING 59 59 ATP (BY SIMILARITY).
FT ACT_SITE 149 149 BY SIMILARITY.
FT DOMAIN 373 378 POLY-GLU.
FT CONFLICT 222 222 P -> R (IN REF. 1).
FT CONFLICT 312 312 V -> M (IN REF. 1).
SQ SEQUENCE 487 AA; 150758EBC5F77D5C CRC64;

Query Match 10.3%; Score 715.5; DB 1; Length 487;
Best Local Similarity 34.0%; Pred. No. 9.4e-16;
Matches 168; Conservative 98; Mismatches 155; Indels 73; Gaps 14;

QY 6 PARSLDDILSAL-RDPAGIFELVVVGNGYGVYKGRVKTGQLAAIKMVDTEDEE 64
Db 10 PRRLKKLDEDSLTKQEEVDFVLEKLEGESYGVYKAIHKETQIIVAIKQVPVESLQ 69
QY 65 EIKOEINMLKKYSHRNATYYGAFIKKSPPGNDQDLWLVMEFCGAGSVTDLVNKTGNA 124
Db 70 IIEISIMQQCDSPH--VVKYGSYFKNT-----DLWIVMEYCGAGSVSDIIR-LRNKT 120
QY 125 LKEDCIAIYCREILRGUHLAHLAHKVIHRDIKQNVLTENAEVKLVDFGVSQLDRTVGR 184
Db 121 LTEDEIATILQSTLKGLEYLHFMKIHRIKAGNILLNTEGHAKLADFGVAGQLTDTMAK 180
QY 185 RNTFIGTPYMAPEVIACDENPDATYDRSDINSLGITALEMARGAPLCDMHPMARFL 244
Db 181 RNTVIGTPYMAPEVI-----QEGTNCVADINSLGITALEMARGKPPYADIHPMRAIFM 235
QY 245 IPRNPPRL-KSKKWSKKFIDFIDTCLIKTYLSRPTEQLLKFPFIRDQPTERQVRIQLK 303
Db 236 IPTNPPPTFRKPELWSDNFTDFVKQCLVKSPEQRATATQLLQHPFVR--SAKGVSI-LR 291
QY 304 DHIDRSRKRGEKEETEYEGSGSEEDDSDHGESEGPSSIMNVPQ----- 347
Db 292 DLINAEADVLRKQESQOREVDQDDEENSEDESDSGTMVRAVGDEMGTVRVASTMTDGA 351
QY 348 -----ESTLRREF--LRLLQENKNSSEALKQCOOLQOQ-----QQRDPBAHKH- 389
Db 352 NTMIEHDTLPSQGLTWVINADEEEEGTWRKRDDETQPAKPSLFYFEQKEKENQNSF 411
QY 390 -----LLHQRRRIEKEKE-----ERRRYEEQQRR-----ERQRKLQEKQRRRL 430
Db 412 GKSVPGLKNSDWDKI PQDGDYEFKSWTVDLQKRLALDPPMMEQIEEIRQYQSKRQ 471
QY 431 EDQALRRREERRQ 444
Db 472 PILDAIEAKRRQ 485

RESULT 11
STK3_HUMAN STANDARD; PRT; 491 AA.
ID STK3_HUMAN
AC Q13188; Q15445; Q15801;

```

16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine protein kinase 3 (EC 2.7.1.37) (STE20-like kinase
MST2) (MST-2) (Mammalian STE20-like protein kinase 2)
DE (Serine/threonine protein kinase Krs-1).
GN STK3 OR MST2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96144292; PubMed=8566796;
RA Creasy C.L., Chernoff J.;
RT "Cloning and characterization of a member of the MST subfamily of
STe20-like kinases.";
RL Gene 167:303-306(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96413604; PubMed=8816758;
RA Taylor L.K., Wang H.C., Erikson R.L.;
RT "Newly identified stress-responsive protein kinases, Krs-1 and Krs-
2.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:10099-10104(1996).
RN [3]
RP SEQUENCE OF 96-203 FROM N.A.
RX MEDLINE=94100173; PubMed=8274451;
RA Schultz S.J., Nigg E.A.;
RT "Identification of 21 novel human protein kinases, including 3 members
of a family related to the cell cycle regulator nima of Aspergillus
nidulans.";
RL Cell Growth Differ. 4:821-830(1993).
CC -!- FUNCTION: OXIDANT STRESS-ACTIVATED SERINE/THREONINE KINASE THAT
MAY PLAY A ROLE IN THE RESPONSE TO ENVIRONMENTAL STRESS (BY
SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN ADULT KIDNEY,
SKELETAL AND PLACENTA TISSUES AND AT VERY LOW LEVELS IN ADULT
HEART, LUNG AND BRAIN TISSUES.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
STE20 subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U26424; AAC50386.1; -;
DR EMBL; U60206; AAB17261.1; -;
DR EMBL; 225422; CAAB0909.1; -;
DR PIR; I38212; I38212.
DR HSSP; P24941; 1HCL.
DR Genew; HGNC:11406; STK3.
DR MIM; 605030; -;
DR GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser thr pkin AS.
DR InterPro; IPR002290; Ser thr_pkinase.
DR Pfam; PF00069; pkinase; I.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding.
DOMAIN 27 278 PROTEIN KINASE.

FT NP_BIND 33 41 ATP (BY SIMILARITY).
FT BINDING 56 56 ATP (BY SIMILARITY).
FT ACT_SITE 146 146 BY SIMILARITY.
FT DOMAIN 308 314 POLY-GLU.
FT DOMAIN 370 375 POLY-GLU.
FT CONFLICT 96 98 WIV -> YLY (IN REF. 3).
FT CONFLICT 121 121 D -> Y (IN REF. 3).
FT CONFLICT 203 203 D -> G (IN REF. 3).
FT CONFLICT 303 303 D -> E (IN REF. 2).
FT CONFLICT 332 334 GEC -> ESV (IN REF. 2).
SQ SEQUENCE 491 AA; 56261 MW; 9CA3B0644F3C14A9 CRC64;
Query Match 10.08; Score 694.5; DB 1; Length 491;
Best Local Similarity 33.6%; Pred. No. 4.1e-15;
Matches 173; Conservative 98; Mismatches 183; Indels 61; Gaps 16;
QY 4 PAPARSLDDLSAL-RDPAGIFELVEVVGNGTYGVYKGRHVKTGLAAIKVMVDTEDE 62
DB 5 PAKSKLKLSEDSLTQKPEEVPDVLKLGSGSYGVFAIKHESGVVAIKQVPVESDL 64
QY 63 EEEIKQEIINLMKKYSHRNATYYGAFIKSPGNDQDLWLVMEFCGAGSVTDLVNKTG 122
DB 65 QEIIK-BISIMQ-CDSPPYVVKYGSYFKNT-----DLWIVMEYCGAGSVSDIIR-LRN 115
QY 123 NALKEDCIATICREILRGLAHLHAHKVTHRDIKGQNVLLTENAELKLVDFGVSAQLDRTV 182
DB 116 KTLIEDIATILKSTLKGLEYLHPMKIHRDIKAGNILLNTEGHAKLADFGVAGQLTDTM 175
QY 183 GRRNFTGTPTWMAPEVIACDENPDATYDRSDIWSLGTITAIENAEAGAPPLCDMHPREAL 242
DB 176 AKRNVITGTFWMAPEVI-----QEIGYNCVADLWSLGTISIEAEGKPPYADHPMEAI 230
QY 243 FLIPRNPPLRL-KSKKWSKFPIDFTCLIKTYLSRPTEQLLPFFI-----RDQ 292
DB 231 FMIPTNPPTFRKPELWSDDFDTEVKKLVKNPQRATATQLQHPFIKNAKPVSIIRD 290
QY 293 PTERQVRIQLKDHDRSKRKEETEYEVSGSEEDDSH-----GEES-EPSSIM 343
DB 291 ITE-AMIKAKRHDEQRELEEEESND-----EDELDSHTMVKTSVGEGETWRTSTM 343
QY 344 NVPGESTLRREFLRLOENKS---NSEALKQQOQQOQQOQDRDPEAHIKHLLHQRRLEE 400
DB 344 SEGATWIEHNSTWLESDLTGTVINSEDEEDGTMRKNTSPQVQRPSPFMDY----PDK 399
QY 401 QKEERRVBEQRRERQRKLEK-----EQORLEDMQALRRERERQAREQRYK 452
DB 400 QDFKXKSHENCNQNHPFPMSKNVFPDNNKVPQDGDGDFLKNLSLELQMLKALDPM 459
QY 453 RKQLEERQSERLQRLQEHAYLKSLOQQOQQOQ 487
DB 460 EREIEELRQRYTAKRQ-----PILDAMDAKKRQO 489
RESULT 12
HPO_DROME
ID_HPO_DROME STANDARD; PRT; 669 AA.
AC Q8T0S6; Q9V8W4;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Serine/threonine protein kinase hippo (EC 2.7.1.37).
GN HPO OR CG11228.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Metopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.; FUNCTION, TISSUE SPECIFICITY, AND INTERACTION WITH
RP SAV AND WTS.
RX MEDLINE=22841274; PubMed=12941274;
RA Harvey K.P., Pfeleger C.M., Hariharan I.K.;
RT "The Drosophila Mst ortholog, hippo, restricts growth and cell

RT proliferation and promotes apoptosis.";
RL Cell 114:457-467(2003).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Berkeley;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frannk C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
RA de Fabios B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupezki M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissensbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.N., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [3]
RP REVISIONS.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0080.22(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley; TISSUE=Head;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource."
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
RN [5]
RP TISSUE SPECIFICITY.
RX MEDLINE=21898525; PubMed=11900973;
RA Brody T., Stivers C., Nagle J., Odenwald W.F.;
RT "Identification of novel Drosophila neural precursor genes using a
RT differential embryonic head cDNA screen."
RL Mech. Dev. 113:41-59(2002).
RN [6]
RP INDUCTION.
RX MEDLINE=22085919; PubMed=12091301;
RA Egger B., Leemans R., Loop T., Kammermeier L., Pan Y., Radimerski T.,
RA Strahm M.C., Certa U., Reichert H.;
RT "Glucogenesis in Drosophila: genome-wide analysis of downstream genes
RT of glial cells missing in the embryonic nervous system."
RL Development 129:3295-3309(2002).
RN [7]
RP FUNCTION, PHOSPHORYLATION OF SAV AND WTS, MUTANT HPO42-47, AND
RP MUTAGENESIS OF LYS-71.
RX MEDLINE=22822128; PubMed=12941273;
RA Wu S., Huang J., Dong J., Pan D.;
RT "hippo encodes a Ste-20 family protein kinase that restricts cell
RT proliferation and promotes apoptosis in conjunction with Salvador and
RT warts".
RL Cell 114:445-456(2003).
CC -I- INDUCTION: Serine/threonine kinase required to restrict cell
CC proliferation and promote apoptosis. Phosphorylates Sav and Wts
CC proteins. The interaction with Sav promotes Wts phosphorylation.
CC Probably part of a same tumor suppression pathway with Sav and
CC Wts. Regulates the level of Th/DAPI apoptosis inhibitor.
CC -I- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -I- TISSUE SPECIFICITY: Expressed in CNS during embryogenesis. In
CC third instar larvae, it is expressed throughout all imaginal
CC disks.
CC -I- INDUCTION: Transcriptionally regulated by Gcm (Glial cells
CC missing).
CC -I- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch.
DR EMBL; AE003795; AAF57543.2; -;
DR EMBL; AY069088; AAL39233.1; -;
DR HSSP; P24941; 1HCL
DR Flybase; FBgn0034453; hpo.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase, 1.
DR ProDom; PD000001; Prot_kinase, 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR PROSITE; PS50013; PROTEIN_KINASE_DOM; 1.
DR Transferase; Serine/threonine-protein kinase; Apoptosis; ATP-binding.
FT DOMAIN 42 293 PROTEIN KINASE.
FT NP BIND 48 56 ATP (BY SIMILARITY).
FT BINDING 71 71 ATP (BY SIMILARITY).
FT ACT SITE 161 161 BY SIMILARITY.
FT DOMAIN 490 503 POLY-GLN.
FT DOMAIN 538 545 POLY-GLN.
FT DOMAIN 490 599 GLN-RICH.
FT MUTAGEN 71 71 K->R: ABOLISHES PHOSPHORYLATION OF SAV.
FT MUTAGEN 166 171 MISSING: IN HPO42-47; NULL MUTANT.
FT CONFLICT 503 503 Q -> CQQQ (IN REF. 1).
SQ SEQUENCE 669 AA; 75109 MW; B3DDDFAP5F4E4EB3 CRC64;
Query Match 10.0%; Score 693.5; DB 1; Length 669;
Best Local Similarity 31.3%; Pred. No. 5.6e-15;
Matches 193; Conservative 104; Mismatches 203; Indels 117; Gaps 22;
Oy 16 SALRDPAGIELVEVNGVGTGYGVYKGRHVKTGOLAIAKVMVDTEDEEEETKQENMLKK 75
Db 33 SLQPPKPVDMYKLGEGSYGVYKAVHKESSIVAIKLVFVSDLHLHIK-EISIMQQ 91


```
DR InterPro; IPR008271; Ser thr pkin_AS.
DR InterPro; IPR002290; Ser thr pkinase.
DR InterPro; IPR001245; Tyr pkinase.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
KW ATP-binding; Transferase; Serine/threonine-protein kinase;
KW Alternative splicing.
FT DOMAIN 20 277
FT BIND 26 34
FT BIND 49 49
FT ACT_SITE 140 140
FT DOMAIN 501 827
FT VARSPIC 294 312
FT CONFLICT 79 79
FT CONFLICT 103 103
FT CONFLICT 118 118
FT CONFLICT 441 441
FT CONFLICT 727 727
SQ SEQUENCE 847 AA; 95044 MW; 3770E0FFC6EECF3 CRC64;
Query Match 10.0%; Score 691; DB 1; Length 847;
Best Local Similarity 36.0%; Pred. No. 8.1e-15;
Matches 169; Conservative 85; Mismatches 149; Indels 66; Gaps 15;
QY 5 APARSLDIDLSALDPAGIFELVWGVNGTYGVYKGRHVNTGQLAAIKWMDVTE-DEE 63
DB 3 APLRPAADI---LRNPQHDYELVORVSGTYGVYKARNVHTGLAAVKKIKLEPGDDF 59
QY 64 EBIQEIENMLKKYSHRNATYYGAFIKKSPGNDQLWLWMEFCGAGSVTDLVKNYKGN 123
DB 60 SLIQEIEFVWKE-CKHCNIVAYFGSLR-----EKLWICMEYCGGSLQDIYHVT--G 110
QY 124 ALKEDCIAYICEILRLGLAHKVIHRDIKQNVILTENAVKLVDPGSAQLDRTVG 183
DB 111 PUSEQIAYVCRETIQLGLAYLTKGMRDRIKGANILLTDHGDVKLADFGVAAKITATIA 170
QY 184 RNTFTIGPYWMAPEVIACDENPDATYDYSRISLIGLITALEMARGAPLCLDMHPMRALF 243
DB 171 KEKSFIGPYWMAPEVAEVEKN--GGYNQLCDIWAAGITALELQELPMPDLHPMRALF 228
QY 244 LIPRN--PPRLUKS-KWKKFIDFTDLTKLYLSRPPTEQLLKFPPIRQPTQVRV 300
DB 229 LMSKSNFQPKLKDRTKWSSTFHNFKVIALTKNPKKPTAEKLLTHTFVGQPLSRALAV 288
QY 301 QLKDHIDRSKRGKEETEYYSSEEDSDHGECEPSSIMNVPGESTLRREFLRLOQ 360
DB 289 ELLDKV-----SNPDHAPYSGDEDD-----LEPHAIIRHTIRSTNRN-----SR 329
QY 361 ENKNSSEALKKOOQQLQOQQOQDPEAHIK-----HLHL-----QRRRIE 399
DB 330 AERTASEINFQKLEFPEPLRKETEAREDEMSLSEENFILWNPFVGDGANTGRTSKRAIP 389
QY 400 EQKEERRRVEQQRRESRQKLEKEQQ---RRLEDMQALRRREERRQA 445
DB 390 PPLPKPRVNTY-----PEDSLPDEKSTIKRCPDLEARAPQVLLRQSS 433
RESULT 14
ID M4K5 HUMAN STANDARD; PRT; 846 AA.
AC Q9Y4K4; Q8IYF6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 5 (EC 2.7.1.37)
```

```
DE DE
GN GN
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, TISSUE
RP SPECIFICITY, AND MUTAGENESIS OF LYS-49.
RC TISSUE=T-cell;
RX MEDLINE=97190217; PubMed=9038372;
RT "A novel human SPS1/STE20 homologue, KHS, activates jun N-terminal
RT kinase.";
RL Oncogene 14:653-659 (1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bonak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.B.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP INTERACTION WITH CRK AND CRKL.
RX MEDLINE=99002481; PubMed=9788432;
RA Oehrl W., Kardinal C., Ruf S., Adermann K., Groffen J., Feng G.-S.,
RA Blenis J., Tan T.-H., Feller S.M.;
RT "The germinal center kinase (GCK)-related protein kinases HPK1 and KHS
RT are candidates for highly selective signal transducers of Crk family
RT adapter proteins.";
RL Oncogene 17:1893-1901 (1998).
CC -|- FUNCTION: May play a role in the response to environmental stress.
CC -|- Appears to act upstream of the c-jun N-terminal pathway.
CC -|- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -|- COPACTOR: Magnesium.
CC -|- SUBUNIT: Interacts with both SH3 domains of the adapter proteins
CC CRK and CRKL.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- TISSUE SPECIFICITY: Ubiquitously expressed in all tissues
CC examined, with high levels in the ovary, testis and prostate.
CC -|- SIMILARITY: Belongs to the Ser/Thr family of protein kinases.
CC STE20 subfamily.
CC -|- SIMILARITY: Contains 1 CNH domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U77129; AB48435.1; -.
EMBL; BC036013; AB36013.1; -.
HSPF; P24941; 1B38.
```



```

CC STE20 subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; X99325; CAAG7700.1; -
DR EMBL; D63780; BAA20420.1; -
DR EMBL; BC007852; RAH07852.1; -
DR PIR; S71886; S71886.
DR HSP; P12931; LEWK.
DR Genew; HGNC:11404; STK25.
DR MIM; 602255; -
DR GO; GO:0004672; F:protein kinase activity; TAS.
DR GO; GO:0006979; P:response to oxidative stress; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser Thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR0109; TYRKINASE.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TKC_1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation.
FT DOMAIN 20 270 PROTEIN KINASE.
FT NP_BIND 26 34 ATP (BY SIMILARITY).
FT BINDING 49 49 ATP (BY SIMILARITY).
FT ACT_SITE 140 140 BY SIMILARITY.
FT CONFLICT 347 348 EP -> DA (IN REF. 1).
SQ SEQUENCE 426 AA; 48111 MW; 183CE5700FCEA716 CRC64;

Query Match 9.8%; Score 681.5; DB 1; Length 426;
Best Local Similarity 38.3%; Pred. No. 8.9e-15;
Matches 169; Conservative 67; Mismatches 124; Indels 81; Gaps 15;

QY 20 DPAGIFELVEVNGTYGVYKGRHVKTGOLAAIKVMDV--TEDEERIKOEINMLKKYS 77
DB 15 DEELFTKLDRIKGSFGEVYKGINHTYKVVVAIKIIDEAEDEIEDIQEITVLSQ-C 73

QY 78 HHRNIATYTGAFIKKSPGNDQDLVMEFCGAGSVTDLVKNTKGNALKEDCIAYICREI 137
DB 74 DSPYITRYFGSYLKT-----KLWIMEYLGGSALDIL--XPGPLEETVIATILREI 124

QY 138 LGLAHLAHAKVHHDIKQNVLLTENAIEVKLVDFGVAQLDRTVGRNRTFTGTYWAP 197
DB 125 LKGLDYLHSEKIHRIKDIKAANVLLSEQGDVKLADFGVAGQLTDTQIKENTFTVGFPMWAP 184

QY 198 EVIACDENPDATYDYSRDIWSLIGITAIEMAGCAPLDCMHPMALFLIPRNPPLKSKK 257
DB 185 EVI-----KQSYDFKALWSLIGITAIELAKGPPNSDLHPMRVFLIPKNSPPTLEGQH 239

QY 258 WSKFIDITDCLIKTYLSRPTEQLKFPFIROQTERQVRIQLKDHIDRSKRKGRKE 317
DB 240 -SKPKFEVEACLNKPRFPPTAKELLKHKFTITRYTKTSFTEL---IDRYKR---WKS 292

QY 318 ETEYISGESEED-DSHGEGEPSPIMNVPGESTLRREFRLIQENKSNSEALKQOQOQLQ 376
DB 293 EGHGESSESDIDGAEADGEGQPIWTFP--PTIR----- 326

QY 377 QQQQRDPEAHIKHLHQRIIEQERRRRVEEQRREREQ-----RKLQEKQ 426
DB 327 -----PSFHSK--LHKGTALHSSQKP-----AEPVKQPRSQCLSTLVRPVFGLKHK 374

QY 427 Q-----RRLEDQALRRE 439

```

Db 375 QSGGSGVGALELENAFSLAE 395

Search completed: August 28, 2004, 00:57:36
Job time : 38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2004, 00:54:25 ; Search time 35 Seconds
(without alignments)
1935.236 Million cell updates/sec

Title: US-10-029-115-2
Perfect score: 6929
Sequence: 1 MGDPAAPASLDIDLSALRD.....SGSSQVYFMTLNRNCIMNW 1312

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A-COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B-COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A-COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B-COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6659	96.1	1326	4	US-09-688-188B-15
2	6659	96.1	1326	4	US-09-291-417D-15
3	4370.5	63.1	1324	4	US-09-645-456A-13
4	4370.5	63.1	1324	4	US-09-425-324A-13
5	4370.5	63.1	1324	4	US-09-645-791-13
6	4360.5	62.9	1332	4	US-09-645-456A-9
7	4360.5	62.9	1332	4	US-09-425-324A-9
8	4360.5	62.9	1332	4	US-09-645-791-9
9	4346	62.7	1353	4	US-09-645-456A-11
10	4346	62.7	1353	4	US-09-425-324A-11
11	4346	62.7	1353	4	US-09-645-791-11
12	4336	62.6	1360	3	US-09-393-569-2
13	4336	62.6	1360	4	US-09-579-664B-14
14	4336	62.6	1360	4	US-09-645-456A-34
15	4336	62.6	1360	4	US-09-425-324A-34
16	4336	62.6	1360	4	US-09-645-791-34
17	4274	61.7	1269	4	US-09-645-456A-15
18	4274	61.7	1269	4	US-09-425-324A-15
19	4274	61.7	1269	4	US-09-645-791-15
20	4264	61.5	1277	4	US-09-645-456A-12
21	4264	61.5	1277	4	US-09-425-324A-12
22	4264	61.5	1277	4	US-09-645-791-12
23	4249.5	61.3	1298	4	US-09-645-456A-14
24	4249.5	61.3	1298	4	US-09-425-324A-14
25	4249.5	61.3	1298	4	US-09-645-791-14
26	4239.5	61.2	1306	4	US-09-645-456A-10
27	4239.5	61.2	1306	4	US-09-425-324A-10

ALIGNMENTS

RESULT 1

US-09-688-188B-15
; Sequence 15, Application US/09688188B
; Patent No. 6656716
; GENERAL INFORMATION:
; APPLICANT: FLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHITE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 039603/0328
; CURRENT APPLICATION NUMBER: US/09/688.188B
; CURRENT FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/291,417
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 1326
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-188B-15

Query Match	96.1%	Score 6659;	DB 4;	Length 1326;
Best Local Similarity	95.8%	Pred. No. 0;		
Matches 1270;	Conservative	3;	Mismatches	3;
Indels	50;	Gaps	3;	
QY	36	YGVYKGRHVKTGQLAAIKVMDVTEDEEEIKQBINMLKKYSHRNIAITYYGAFIKKSP 95		
DB	2	FGEVYEGRHVKTGQLAAIKVMDVTEDEEEIKQBINMLKKYSHRNIAITYYGAFIKKSP 61		
QY	96	GNDQMLVMEFCGAGSVTDLVKNTKGNALKEDECIAYICRILRGLAHLHAHKVHRDIK 155		
DB	62	GNDQMLVMEFCGAGSVTDLVKNTKGNALKEDECIAYICRILRGLAHLHAHKVHRDIK 121		
QY	156	QNVLLTENAEVKLVDFGVSAQLDRTVGRNRTFGTPYVMAPEVTACDENPDATYDVRSD 215		
DB	122	QNVLLTENAEVKLVDFGVSAQLDRTVGRNRTFGTPYVMAPEVTACDENPDATYDVRSD 181		
QY	216	IWSLGITAIEMAEAPPLCDMHPRALFLIPRNPPLPKSKKWKFFIDFIDTCLIKTYL 275		
DB	192	IWSLGITAIEMAEAPPLCDMHPRALFLIPRNPPLPKSKKWKFFIDFIDTCLIKTYL 241		
QY	276	SRPTEQLLKPFPIRDQPTQVRQLKDHIDRSKRGKEKEETEYVSGSEEDDSHGE 335		
DB	242	SRPTEQLLKPFPIRDQPTQVRQLKDHIDRSKRGKEKEETEYVSGSEEDDSHGE 301		
QY	336	EGEPSSINNVGSESTLRREFLRLOENKNSSEALKQOQLOOQOORDPEAHKHLHORQ 395		

302	Qy	EGEPSSIMVPGESTLRRREFLRLOQENKSNSEALKQQOQLQQOQRPDEPAHIKHLHQRC	361
396	Qy	RRIEQKEERRRVVEEQORREREQRKLQEKEOQRLEDQOALRRBEERRQAREQ	449
362	Db	RRIEQKEERRRVVEEQORREREQRKLQEKEOQRLEDQOALRRBEERRQAREQEIYIHR	421
450	Qy	-----EYKXKQLEBEOQSERLQROLOEHAYLKSLOOQOQOQ	486
422	Db	LEEBQRLIILQOQLLOEQALLLEYKXKQLEBEOQSERLQROLOEHAYLKSLOOQOQOQ	481
487	Qy	QLOKQOQOQLLPGDRKPLIYHGRGWNPADKPAWAVEEERTMKNQONSPLAKSPGSGT	546
482	Db	QLOKQOQOQLLPGDRKPLIYHGRGWNPADKPAWAVEEERTMKNQONSPLAKSPGSGT	541
547	Qy	PEPPIPOASPGPPGSLQTPPMQRPVPBQEGPHK-----SLOQOP	586
542	Db	PEPPIPOASPGPPGSLQTPPMQRPVPBQEGPHKSLVAHRVPLKPYAAPVPRSQSLOQOP	601
587	Qy	TRNLAAFPASHDDPPAIPAPTATPSARGAVIRONSDDPTSEGPGSPNPPPAWVRPDNEAPP	646
602	Db	TRNLAAFPASHDDPPAIPAPTATPSARGAVIRONSDDPTSEGPGSPNPPPAWVRPDNEAPP	661
647	Qy	KVPORTSSIALNTSGAGSRPAQAVRARPRSNAMQIYLQRAERGTPKPPGPPAQP	706
662	Db	KVPORTSSIALNTSGAGSRPAQAVRARPRSNAMQIYLQRAERGTPKPPGPPAQP	721
707	Qy	GPPNASSNPDLRRSDPCWERSDSVLPASHGHLPOAGSLERNVRGASKLDSPPVLSFGNK	766
722	Db	GPPNASSNPDLRRSDPCWERSDSVLPASHGHLPOAGSLERNVRGASKLDSPPVLSFGNK	781
767	Qy	AKPDHRSRQGRPADFVLLKERTLEDAAPRPKAMDYSSSEEVESSEDEEEGEGGPAE	826
782	Db	AKPDHRSRQGRPADFVLLKERTLEDAAPRPKAMDYSSSEEVESSEDEEEGEGGPAE	841
827	Qy	GSRTTPGSRSDGDTDSVSVMVVDHVEITGTQPPYGGGTVMVQRTPEERNLLHADSNXY	886
842	Db	GSRTTPGGR-DGDTDSVSVMVVDHVEITGTQPPYGGGTVMVQRTPEERNLLHADSNXY	900
887	Qy	TNLPDVVQPSHSPTENSKGSPSPSKDGGDYQSRGLVKAPCKSFTWFDLGIYQPGSG	946
901	Db	TNLPDVVQPSHSPTENSKGSPSPSKDGGDYQSRGLVKAPCKSFTWFDLGIYQPGSG	960
947	Qy	DSIPITALVGEGTRLQLOQYDVRKGSVMVNPNTTRAHSETPEIRYKXKFNSEILCAA	1006
961	Db	DSIPITALVGEGTRLQLOQYDVRKGSVMVNPNTTRAHSETPEIRYKXKFNSEILCAA	1020
1007	Qy	LWGNVLVGTENGMLLDRSQGKVYGLIGRRRQOMDVL EGLNLLTISGKRNKLAVY	1066
1021	Db	LWGNVLVGTENGMLLDRSQGKVYGLIGRRRQOMDVL EGLNLLTISGKRNKLAVY	1080
1067	Qy	LSWURNKILHNDPVEKKQGTWTVGDMEGCGHYRWKYRIKFLVIALKSSVEVYAWAPK	1126
1081	Db	LSWURNKILHNDPVEKKQGTWTVGDMEGCGHYRWKYRIKFLVIALKSSVEVYAWAPK	1140
1127	Qy	PYHKFMAFKFADLPHRPLLVDLTVBERQRLKVITYGSSAGFAHVDVDSGNSYDIYIPVHI	1186
1141	Db	PYHKFMAFKFADLPHRPLLVDLTVBERQRLKVITYGSSAGFAHVDVDSGNSYDIYIPVHI	1200
1187	Qy	QSQITTHAIIIFLPTDGMEMLLCYEDBGVVYNTGRIIKDVVLQWGMPTSVIYICNSQI	1246
1201	Db	QSQITTHAIIIFLPTDGMEMLLCYEDBGVVYNTGRIIKDVVLQWGMPTSVIYICNSQI	1260
1247	Qy	MWGGEKAIEIRSVETGHLGDGVFMHKAQRLKFLCERNDKVFFASVRSGGSSQVYFMTLNR	1306
1261	Db	MWGGEKAIEIRSVETGHLGDGVFMHKAQRLKFLCERNDKVFFASVRSGGSSQVYFMTLNR	1320
1307	Qy	NCINMW	1312
1321	Db	NRINMW	1326

```
; Sequence 15, Application US/09291417D
; Patent No. 6680170
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0329
; CURRENT APPLICATION NUMBER: US/09/291,417D
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 1326
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-291-417D-15

Query Match          96.1%; Score 6659; DB 4; Length 1326;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 1270; Conservative 3; Mismatches 3; Indels 50; Gaps 3;

Qy      36 YGVYKGRHVTKQLAAIKVMVTEDEEBIKQEINMLKKYSHRNIATYYGAFIKKSPP 95
       :|::|||
Db      2 FGVEYEGRHVKTGOLAAIKVMVTEDEEBIKQEINMLKKYSHRNIATYYGAFIKKSPP 61

Qy      96 GNDDQLWLWMEFCGAGSVTDLVKNTGNALKEDCIAYICREILRGLAHLHAHKVIHRDIK 155
       ||::|||
Db      62 GNDDQLWLWMEFCGAGSVTDLVKNTGNALKEDCIAYICREILRGLAHLHAHKVIHRDIK 121

Qy      156 GQNVLITENAEVKLVDFGVSQAOLDRTVGRRNTEFTIGTPYMAPEVIACDENPDATDYRS 215
       ||::|||
Db      122 GQNVLITENAEVKLVDFGVSQAOLDRTVGRRNTEFTIGTPYMAPEVIACDENPDATDYRS 181

Qy      216 IWSLGIATAEMAAGAPPLCDMHPMRALFLIPRNPPRLKSKWKSKKFIDFIDTCLIKTYL 275
       ||::|||
Db      182 IWSLGIATAEMAAGAPPLCDMHPMRALFLIPRNPPRLKSKWKSKKFIDFIDTCLIKTYL 241

Qy      276 SRPPTQLLKFPPIRDQPTERQVRIQLKDHI DSRKKRGKEKETETEYSGSEEDDSHGE 335
       ||::|||
Db      242 SRPPTQLLKFPPIRDQPTERQVRIQLKDHI DSRKKRGKEKETETEYSGSEEDDSHGE 301

Qy      336 EGPSSIMNVPGESTLRREFLRLOQSNKNSEALKQQOQLQQOQQORDPEAHIKHLLHQ 395
       ||::|||
Db      302 EGPSSIMNVPGESTLRREFLRLOQSNKNSEALKQQOQLQQOQQORDPEAHIKHLLHQ 361

Qy      396 RRIEEQEKERRRYVEEQRREREQRKLQOEQQRRLEDMQALRRBEERRRAERQ----- 449
       ||::|||
Db      362 RRIEEQEKERRRYVEEQRREREQRKLQOEQQRRLEDMQALRRBEERRRAERQEVIRHR 421

Qy      450 -----EYRKQL EEQRQSERLQRLQOE HAYLKS LQQOQQOQQ 486
       ||::|||
Db      422 LEEERQRL ETLQOQLLQE QALLLEYKRQL EEQRQSERLQRLQOE HAYLKS LQQOQQOQQ 481

Qy      487 QLQCKQQOQLLGDRKPLVHYHGMMNPADKPAPAWAVEERTNMNKQNSPLAKSKPGSTG 546
       ||::|||
Db      482 QLQCKQQOQLLGDRKPLVHYHGMMNPADKPAPAWAVEERTNMNKQNSPLAKSKPGSTG 541

Qy      547 PEPPIPQASGGPPGPLSQTPPMQRPVPEQEPGHK-----SLQDQP 586
       ||::|||
Db      542 PEPPIPQASGGPPGPLSQTPPMQRPVPEQEPGHKSLVAHRVLPKYAAPVRPSQS LQDQP 601

Qy      587 TRNLAAFPAASHDDPAIPAPTATPSARGAVIRQNSDPTSEGPGCPSPNPAPWVPDNEAPP 646
       ||::|||
Db      602 TRNLAAFPAASHDDPAIPAPTATPSARGAVIRQNSDPTSEGPGCPSPNPAPWVPDNEAPP 661

Qy      647 KVPQRTSSITALNTSGAGSRPAQAVRAPRPRSNSAWQIYLQRAAREGTPKPPGPPAOPP 706
       ||::|||
Db      662 KVPQRTSSITALNTSGAGSRPAQAVRAPRPRSNSAWQIYLQRAAREGTPKPPGPPAOPP 721

Qy      707 GPVNASSNFLRRSDFGWERSDVLPASHGLLPQAGSLERNRVGASSKLDSSPVLSFGNK 766
```

```
Db 722 GPNASSNPDDRSGRSDVLPASHGHLPPQAGSLRNRVGSXPDSPLSPGNK 781
Qy 767 AKPDHRRSRGPRPADVLLKERTLEAPRPKAMDYSSSEVESSDDDEEGEGGPAE 826
Db 782 AKPDHRRSRGPRPADVLLKERTLEAPRPKAMDYSSSEVESSDDDEEGEGGPAE 841
Qy 827 GSRDTPGGRSGDTSVSTVMVHDVVEITGTQPPYGGGTVMVQRTPEERMLLHADSNY 886
Db 842 GSRDTPGGR--DGDTSVSTVMVHDVVEITGTQPPYGGGTVMVQRTPEERMLLHADSNY 900
Qy 887 TNLDPVQPSHSPTESSKQSPPSKDGSGDVQSRGLVKAPGKSSFTMFVDLGIYQPGSSG 946
Db 901 TNLDPVQPSHSPTESSKQSPPSKDGSGDVQSRGLVKAPGKSSFTMFVDLGIYQPGSSG 960
Qy 947 DSIPITLVGGEGTRLDQLQYDVRKGSVVNNPTNTRAHSETPEIRKYKGFNSBILCAA 1006
Db 961 DSIPITLVGGEGTRLDQLQYDVRKGSVVNNPTNTRAHSETPEIRKYKGFNSBILCAA 1020
Qy 1007 LWGVNLLVGTENGLMLLDRSGGKGYGLIGRRRFQOMDVLGLNLLITISGRNKLRYVY 1066
Db 1021 LWGVNLLVGTENGLMLLDRSGGKGYGLIGRRRFQOMDVLGLNLLITISGRNKLRYVY 1080
Qy 1067 LSWLRNKLHNDPEVEKKGQMTTVGDMGCGHYRVVYKFKFLVIALKSSVEVYAWAPK 1126
Db 1081 LSWLRNKLHNDPEVEKKGQMTTVGDMGCGHYRVVYKFKFLVIALKSSVEVYAWAPK 1140
Qy 1127 PYHKPMAFKSPADLPRLVLLVTLVEEGORLKVYIGSSAGFHAVDVDSGNSDYIYPVHI 1186
Db 1141 PYHKPMAFKSPADLPRLVLLVTLVEEGORLKVYIGSSAGFHAVDVDSGNSDYIYPVHI 1200
Qy 1187 QSQITPHAIIFLPNTDGMMLLCYDEGVYVNTYGRIIKDVVLQWGMPTSVAYICSNQI 1246
Db 1201 QSQITPHAIIFLPNTDGMMLLCYDEGVYVNTYGRIIKDVVLQWGMPTSVAYICSNQI 1260
Qy 1247 MGWGEKATEIRSVETGHLDDGVFMHKAQRLKFLCERNDKVFPASVRSQSSQVYFMTLNR 1306
Db 1261 MGWGEKATEIRSVETGHLDDGVFMHKAQRLKFLCERNDKVFPASVRSQSSQVYFMTLNR 1320
Qy 1307 NCIMNW 1312
Db 1321 NRIMNW 1326

RESULT 3
US-09-645-456A-13
; Sequence 13, Application US/09645456A
; Patent No. 6562580
; GENERAL INFORMATION:
; APPLICANT: FU, C. Alan
; TITLE OF INVENTION: NOVEL GERMINAL CENTER KINASE CELL CYCLE PROTEINS, COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: A-68344/RMS/DHR
; CURRENT APPLICATION NUMBER: US/09/645,456A
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US/09/425,324
; PRIOR FILING DATE: 1999-10-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 1324
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-645-456A-13

Query Match 63.1%; Score 4370.5; DB 4; Length 1324;
Best Local Similarity 64.7%; Pred. No. 7.8e-258;
Matches 899; Conservative 137; Mismatches 210; Indels 143; Gaps 31;

Qy 1 MGDPAARSLLDDIDLSALRDPAGIFELVELVGNCTYGVYKGRHVKTQLAAIKVMDVTG 60
```

```
Db 1 MASDPAARSLLDDIDLSALRDPAGIFELVELVGNCTYGVYKGRHVKTQLAAIKVMDVTG 60
Qy 61 DEEBEIKOEINMLKKYSHRNIIATYYGAFIKKSPGNDQLWLVMFCGAGSVTDLVKQT 120
Db 61 DEEBEIKOEINMLKKYSHRNIIATYYGAFIKKSPGNDQLWLVMFCGAGSVTDLVKQT 120
Qy 121 KGNALKEDCAYICRETLRGLAHKAVHHRDIKQONVLLTENAELVKLVDFGVSAQLDR 180
Db 121 KGNLTKBEWAIYICRETLRGLSHLHQHKVTHRDIKQONVLLTENAELVKLVDFGVSAQLDR 180
Qy 181 TVGRNRTFIGTPYMAPEVIACDENPDATYDYSIDIMSLGITAEMAEGAPPLCDMHPMR 240
Db 181 TVGRNRTFIGTPYMAPEVIACDENPDATYDYSIDIMSLGITAEMAEGAPPLCDMHPMR 240
Qy 241 ALFLIPNPPRLKSKWKKKFDIDTCLIKTVLSRPPTQLKFPFIRDOPTERQVRI 300
Db 241 ALFLIPNPPRLKSKWKKKFSFIESCLVKNHSQRPATEQLMKHPFIRDOPTERQVRI 300
Qy 301 QLKDHIDRSKKGKEETEVEYSGSEEDDSHGEEGEPSSIMNVPGESTLRRREFLRQQ 360
Db 301 QLKDHIDRTKKKGKEDETEYSGSEEBEEN--DSGEPSILNLPGESTLURRDFLRQL 359
Qy 361 ENKNSBALKQOQOQOQOQDPEAHIKHLLHQRRIEEOKEERRRVEEQORREREQRK 420
Db 360 ANKERSBALRQO--LEQOQRENEEHKQLLAERQKIEEQKEQRRLEEQORREKELRK 417
Qy 421 LOEKEQORRLLEDQALREERERQAEQYKRYKQLEEQSERLQRLQOQEHAYLSLQ 480
Db 418 QOEREQRHRYBEQ--MRREERRRAHEQYKRYKQLEEQSERLQRLQOQERDYLVSQ 475
Qy 481 QOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 540
Db 476 HQREQR--PVEKKPLYHYKEGSPSEKPAWAKEVEERSRLNR--QSSPAMPH 524
Qy 541 KPGSTGPEPPIQASPG-----PPGPLSOTPPMQRPVEPO-----EGP-----HKS 581
Db 525 KVANRISDPNLPPESEFSISGVQP---ARTPPMLRPVDPQI PHLVAKSGPALTASQS 581
Qy 582 LQDOPTNLNLAAPASHDPDPAIPAPTATPSARGAVIRONSDDPTSEGPGPSP-----NPP 635
Db 582 VHEQPTKGLSGFQEALN-----VTSHRVEMPRQNSDPTSENP--PLPTRIEKFDRS 630
Qy 636 AWVRPDNEAPPKVPQRTSSITATALTSGAGSRPAQAVRAPRPSNSAWQIYLORRAERG 695
Db 631 SWLRQEEDI PPKEVPQRTTISPALARKNPGNGSALG-----PRIGS----- 672
Qy 696 PKPPGPPAQPDPNPNASSNPDLRRSDPGWE-----RSDSVLPASH-----GHLP- 739
Db 673 -----QPIRASNPDLRTERTEPILSPLORTSSGSSSSSTPSSQPSGGSGQP 720
Qy 740 -QAGSLERNRVGASSKLDSPVL--SPGNKAKPDHR--SRGPRPADFVL-----LKERTL 790
Db 721 SQAGSSERTVRANSKSEGSFVLPHPEA-KVKPEESRDITRPSRPADILATALAKELREL 779
Qy 791 DEAPRPPKAMDYSSSEVESSDEDEEGEGGPAES---RDTP-----GGRSDGDTDS 842
Db 780 BETNRPMKKVTDYSSSESESESESESESESESESESESESESESESESESESESESE 839
Qy 843 VSTMVVDHVEITGTQPPYGG-----GTMVQVQRTPEERMLLHADSNYGT---NLPDVVQ 894
Db 840 VGMVGTGLE--TSHADSPGSGISREGTLMIRETSGEKKRSGHSDSNGFAGHINLPDLVQ 897
Qy 895 PSHS-----PTENSKGQSPPKD--GSGDYQSRGLVKAFKGSFTWFDLGIYQPGS---- 945
Db 898 QSHSPAGTPTTEGLGRVSTHQEMDSGTGYGNG---SSTKASFTFPVDPVPRVYQTSPTDE 954
Qy 946 -GDSIPITLVGGEGTRLDQLQY--DVKPGSVVNNPTNTRAHSETPEIRKYKGFNSBIL 1003
Db 955 EDESSAALFTSELLRQEQAKLNEARKISVVNNPTNTRPHSDTPEIRKYKGFNSBIL 1014
Qy 1004 CAALWGVNLLVGTENGLMLLDRSGGKGYGLIGRRRFQOMDVLGLNLLITISGRNKL 1063
Db 1015 CAALWGVNLLVGTENGLMLLDRSGGKGYGLIGRRRFQOMDVLGLNLLITISGRNKL 1074
```



```
Db      840  SETHDGTAVSDIPLRIPTGAPGSEYQVNVGMVTHGLE--TSHADSPSGSISREGTLM 897
Qy      869  QRTPEERNLLHADSNYGT---NLPDVVQPSHS---PTENSKGQSPPSKD--GSGDYQSR 920
Db      898  RETSGEKRSCHSDSNAGFAGHINLPDLVQSHSPAGTPTGELGRVSTHSGEMDSGTGYGM 957
Qy      921  GLVKAPGKSSFTMFVDLGIYQPGS-----GDSIPITALVGGEGTRLDQLQY--DVRKGSV 974
Db      958  G---SSTKASFTPFVDPRVYQTSPTDEDEBESSAAALFTSELLRQEQAKLNEARKISV 1014
Qy      975  VNVNPTNTRAHSETPEIRKYKRFNSEILCAALWGNLLVGTENGMLLDRSGQGKVGYL 1034
Db      1015  VNVNPTNTRAHSETPEIRKYKRFNSEILCAALWGNLLVGTENGMLLDRSGQGKVGYL 1074
Qy      1035  IGRRFQOMDVLGLNLLITISGKRNKLRYVYLSWLNKI LNDPVEVEKKQGTWTVGDM 1094
Db      1075  INRRFQOMDVLGLNLLITISGKRNKLRYVYLSWLNKI LNDPVEVEKKQGTWTVGDM 1134
Qy      1095  GCGHYRVVKYERIKFVLIALKSVVEYAWAPKPYHKFMFKSPADLPHRPLLDLTVEEG 1154
Db      1135  GCIHYKVKYERIKFVLIALKSVVEYAWAPKPYHKFMFKSPADLPHRPLLDLTVEEG 1194
Qy      1155  QRLKVIYGVSSAGFHAVDVDSGNSYDIYIPVHIQSOITTHAIIFLPTDGMELLCYDEBG 1214
Db      1195  QRLKVIYGVSSAGFHAVDVDSGNSYDIYIPVHIQSOITTHAIIFLPTDGMELLCYDEBG 1254
Qy      1215  VYVNTYGRITKDVVLQWGMPTSVAYICNSQIMGMGEKAIEIRSVETGHLGDFVHMKRAQ 1274
Db      1255  VYVNTYGRITKDVVLQWGMPTSVAYICNSQIMGMGEKAIEIRSVETGHLGDFVHMKRAQ 1314
Qy      1275  RLKFLCERNDKVFFASVRSVSGSSQVYFMTLNRNCIMNW 1312
Db      1315  RLKFLCERNDKVFFASVRSVSGSSQVYFMTLNRNCIMNW 1352

RESULT 10
US-09-425-324A-11
; Sequence 11, Application US/09425324A
; Patent No. 6562591
; GENERAL INFORMATION:
; APPLICANT: Fu, C. Alan
; TITLE OF INVENTION: NOVEL GERMINAL CENTER KINASE CELL CYCLE PROTEINS, COMPOSITIONS AND
; FILE OF INVENTION: METHODS OF USE
; FILE REFERENCE: A-68344/RMS/DRH
; CURRENT APPLICATION NUMBER: US/09/425,324A
; CURRENT FILING DATE: 1999-10-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 1353
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-425-324A-11

Query Match      62.7%; Score 4346; DB 4; Length 1353;
Best Local Similarity 63.4%; Pred. No. 2.5e-256;
Matches 899; Conservative 137; Mismatches 210; Indels 172; Gaps 32;

Qy      1  MGDPAARSLDDIDLSALRDPAGIFELVEVWGNGTYGVYQGRHVKTGQLAAIKVMDYTE 60
Db      1  MASDPAARSLDDIDLSALRDPAGIFELVELVNGTYGVYQGRHVKTGQLAAIKVMDVTG 60
Qy      61  DEEBEIKOBINMLKKYSHRNATYYGAFIKSPGNDOLWLVMEFCGAGSVTLVKN 120
Db      61  DEEBEIKOBINMLKKYSHRNATYYGAFIKKPNPMDQLWLVMEFCGAGSVTLVKN 120
Qy      121  KGNALKEDCIAYICREILAGLAHLAHKVIHRDIKQGNVLLTENAELVDFGVSAQLDR 180
Db      121  KGNLKEEWIAYICREILAGLSHLHQHKVIHRDIKQGNVLLTENAELVDFGVSAQLDR 180
```

```
Qy      181  TVGRENTEIGTPYMAPEVIACDENPDATYDYSRDIWSLIGITATEMASGAPFLCDHMPMR 240
Db      181  TVGRENTEIGTPYMAPEVIACDENPDATYDFKSLDWSLIGITATEMASGAPFLCDHMPMR 240
Qy      241  ALFLIPRNPRLKSKWKKKFFIDFTCLIKTYLSRPTEQLLKFPPIRQDPTERQVRI 300
Db      241  ALFLIPRNPRLKSKWKKKFFSQSFIESCLVKNHSQRPATEQLMKHPPIRQDPTERQVRI 300
Qy      301  QLKDHIDSRKKREKETEYEGSBEEDSDHSEEGEPSSIMVPGESTLREPLRLQO 360
Db      301  QLKDHIDTRKKRGEKETEYEGSBEEBEEN--DSGEFSSILNLPGETSLRRDLRLQL 359
Qy      361  ENKNSSEALKQOQQLQOQOQDPEAHIKHLHORRRIEEOKEERRRVEEQORRREROK 420
Db      360  ANKERSEALRQO--LEOQORENEEHKQQLAERQKRIEEOKEORRRLEEQORREKELRK 417
Qy      421  LQKEEQORRLDMOALRREERROAERQ-----EY 451
Db      418  QQEREQRHVEEQ--MRREERRRAEHQEVIRQLBEEQRQLEILQOQLLHEQALLLEY 475
Qy      452  KRQLEEQORERLORLOQOEHAYLKSLOOQOQOQLOQOQOQQLLPGDRKPLHYHGRM 511
Db      476  KRQLEEQORERLORLOQOERDYLVSLOHQOERQ-----PVEKKPLHYHKEGM 525
Qy      512  NPADKPAWAREVEERTMKNQONSPLAKSPGSGTGPPEPIPOASPG-----PGPLSOT 565
Db      526  SPSEKPAWAKEVEBERSRLNR--QSSPAMPHKIVANRISDNLPSPRESFISGVQP---ART 581
Qy      566  PPMORPVEPQ-----EGP-----HKSLOQOQTRNLAAFPASHDDPPAIPAPTATPSA 612
Db      582  PPMLRPVDVQPLPHLVAVKSGQPALTASQSVHEQPTKGLSGFOEALN-----VTSH 631
Qy      613  RGAVIRQNSDPTSGPGSP-----NPPAWVRPDNEAPPKVQORTSIAIATNTSGAGG 666
Db      632  RVEMPRQNSDPTSENP--PLPTRIEKFDRSSWLRQEBDIPPKVPQORTTSISPALARKNSPG 690
Qy      667  SRPAQAVPARPRSNSAQIYLQRAERCTPKPPGPPAPPPNASSNPDLRRSDPGWE- 725
Db      691  NGSALG-----PRLGS-----QPIRASNPDLRRTEPILES 720
Qy      726  -----RSDSVLPASH-----GHLP--QAGSLERNRVGASSKLDSSPVL--SPGNKA 767
Db      721  PLQRTSSGSSSSSTPSSQSGSQSGSQSGQSGSERTRVANSKSEGSVLPHEPA-KV 779
Qy      768  KPDDHR--SRFGRPADFVL-----LKERTLDRAAPPKAMDYSSSESESESEDEEERGE 821
Db      780  KPESRDITRPSRPADLTALAKELRELRIETNRPMKKVTDYSSSESESESESESEDEE 839
Qy      822  GGPAEGS---RDTF-----GGRSDGDTDSVSTMVVDVEITGTQPPYGG-----GTWV 868
Db      840  SETHDGTAVASDIPLRIPTGAPGSEYQVNVGMVTHGLE--TSHADSPSGSISREGTLM 897
Qy      869  QRTPEERNLLHADSNYGT---NLPDVVQPSHS---PTENSKGQSPPSKD--GSGDYQSR 920
Db      898  RETSGEKRSCHSDSNAGFAGHINLPDLVQSHSPAGTPTGELGRVSTHSGEMDSGTGYGM 957
Qy      921  GLVKAPGKSSFTMFVDLGIYQPGS-----GDSIPITALVGGEGTRLDQLQY--DVRKGSV 974
Db      958  G---SSTKASFTPFVDPRVYQTSPTDEDEBESSAAALFTSELLRQEQAKLNEARKISV 1014
Qy      975  VNVNPTNTRAHSETPEIRKYKRFNSEILCAALWGNLLVGTENGMLLDRSGQGKVGYL 1034
Db      1015  VNVNPTNTRAHSETPEIRKYKRFNSEILCAALWGNLLVGTENGMLLDRSGQGKVGYL 1074
Qy      1035  IGRRFQOMDVLGLNLLITISGKRNKLRYVYLSWLNKI LNDPVEVEKKQGTWTVGDM 1094
Db      1075  INRRFQOMDVLGLNLLITISGKRNKLRYVYLSWLNKI LNDPVEVEKKQGTWTVGDM 1134
Qy      1095  GCGHYRVVKYERIKFVLIALKSVVEYAWAPKPYHKFMFKSPADLPHRPLLDLTVEEG 1154
Db      1135  GCIHYKVKYERIKFVLIALKSVVEYAWAPKPYHKFMFKSPADLPHRPLLDLTVEEG 1194
Qy      1155  QRLKVIYGVSSAGFHAVDVDSGNSYDIYIPVHIQSOITTHAIIFLPTDGMELLCYDEBG 1214
```

```

Db 1195 QRLKVFEGSHGTFPVIDVDGNSYDIYPISHIQGNITPAHIVLPKTDGKEMLVCYVEDEG 1254
Qy 1215 VYVNTYGRIIKDVVLQWGEPTSVAYICSNQIMGWGEKATEIRSVETGHLDDGVFMHKRAQ 1274
Db 1255 VYVNTYGRITKDVVLQWGEPTSVAYIHSNQIMGWGEKATEIRSVETGHLDDGVFMHKRAQ 1314
Qy 1275 RLKFLCERNDKVFFASVRSQSSQVYFMTLNRNCIMNW 1312
Db 1315 RLKFLCERNDKVFFASVRSQSSQVYFMTLNRNSMWNW 1352

RESULT 11
US-09-645-791-11
: Sequence 11, Application US/09645791
: Patent No. 659658
: GENERAL INFORMATION:
: APPLICANT: Luo, Ying
: APPLICANT: Fu, Alan C
: APPLICANT: Shen, Mary
: TITLE OF INVENTION: NOVEL GERMINAL CENTER KINASE CELL CYCLE PROTEINS; COMPOSITIONS AND
: FILE REFERENCE: A-68344-1/RMS/DHR
: CURRENT APPLICATION NUMBER: US/09/645,791
: CURRENT FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 09/425,324
: PRIOR FILING DATE: 1999-10-21
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 11
: LENGTH: 1353
: TYPE: PRT
: ORGANISM: Artificial sequence
: FEATURE:
: OTHER INFORMATION: synthetic
US-09-645-791-11

Query Match 62.7%; Score 4346; DB 4; Length 1353;
Best Local Similarity 63.4%; Pred. No. 2,5e-256;
Matches 899; Conservative 137; Mismatches 210; Indels 172; Gaps 32;

Qy 1 MGDPAPARSDDIDLALRDPAGIFELVGVNGTYGVYKGRHVKTGLAAIKVMDVTE 60
Db 1 MASDPSARSLDEIDLALRDPAGIFELVELVNGTYGVYKGRHVKTGLAAIKVMDVTG 60
Qy 61 DEEEIKQEIINMLKYSHHENIATYGAFTKSPGNDQDLWLVMEFCGAGSVTDLVKNT 120
Db 61 DEEEIKQEIINMLKYSHHENIATYGAFTKKNPGMDQDLWLVMEFCGAGSVTDLIKNT 120
Qy 121 KGNALKEDCIAYICREILRGLAHLAHKVIHROI KGNVLLTENAELVKLVDFGVSAQLDR 180
Db 121 KGNLTKEEWAYICREILRGLSHLHKVHROI KGNVLLTENAELVKLVDFGVSAQLDR 180
Qy 181 TVGRNFTIGTPYMAPEVIACDENPDATYDYSIDWSLGITAIEMAGAPPLCDMHPMR 240
Db 181 TVGRNFTIGTPYMAPEVIACDENPDATYDFKSDLWSLGITAIEMAGAPPLCDMHPMR 240
Qy 241 ALFLIPNPPRLKSKWKKFDFDITCLIKTYLSRPPTQQLKPPFIRDQPTERQVRI 300
Db 241 ALFLIPNAPRLKSKWKKFDFESCIIVKHSQRPAEQQLMKHPFIRDQPTNERQVRI 300
Qy 301 QLKDHIDRSKKGKEKETETEVYSGSEEDDSHGEGEPSSIMNVPGESTLRREFLRLOQ 360
Db 301 QLKDHIDRTKKKGKEDTETEVYSGSEEEEN-DSGEPSSIINLPGESTLRDRFLRLQL 359
Qy 361 ENKSNSEALKQQOQLQQOQQORDPEAHIKHLHQRRRIEQQKEERRRVEQQORREQRK 420
Db 360 ANKERSALRRQQ--LFOQQKRENEHKRQLLAERQKRIEBOKRRRLERQQORREKELR 417
Qy 421 LOEKEQORRLDQWALRRERERQAEREQ-----EY 451
Db 418 QQEREQRRHYEEQ--MREEREERRAEHQEYIRBQLEEQRIEQLQQQLLHEQALLEY 475

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2004, 00:43:18 ; Search time 154 Seconds
(without alignments)
2407.158 Million cell updates/sec

Title: US-10-029-115-2
Perfect score: 6929
Sequence: 1 MGDPAFSLDIDLSALRD.....SGSSQVYFMTLNRNCINW 1312

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6929	100.0	1312	7 ADE34150	Ade34150 Human mis
2	6660.5	96.1	1295	5 AAO18508	AAO18508 Human ins
3	6659	96.1	1326	2 AAY55933	Aay55933 Human ZC3
4	6646.5	95.9	1303	4 AAM79153	Aam79153 Human pro
5	6646.5	95.9	1303	5 AAO18507	AAO18507 Human ins
6	6628.5	95.7	1303	4 AAE10612	AAE10612 Human nov
7	6622	95.6	1332	4 AAE10611	AAE10611 Human nov
8	6603.5	95.3	1276	7 ADE34154	Ade34154 Human mis
9	6311	91.1	1244	4 AAE10614	AAE10614 Human nov
10	6304.5	91.0	1273	4 AAE10613	AAE10613 Human nov
11	5651	81.6	1701	4 ABG24020	Abg24020 Novel hum
12	4370.5	63.1	1324	4 AAB68221	Aab68221 Amino aci
13	4360.5	62.9	1332	4 AAB68217	Aab68217 Amino aci
14	4346	62.7	1353	4 AAB68219	Aab68219 Amino aci
15	4336	62.6	1360	3 AAY85263	Aay85263 Human pro
16	4336	62.6	1360	4 AAB50059	Aab50059 Large NIK
17	4336	62.6	1360	5 AAO18506	AAO18506 Human ins
18	4336	62.6	1385	7 ADP14162	Adp14162 Human sfc
19	4325	62.4	1385	4 AAM79405	Aam79405 Human pro
20	4274	61.7	1268	7 ADE31765	Ade31765 Human 154
21	4274	61.7	1269	4 AAB68223	Aab68223 Amino aci
22	4264	61.5	1277	4 AAB68220	Aab68220 Amino aci
23	4249.5	61.3	1298	4 AAB68222	Aab68222 Amino aci
24	4244	61.2	1339	4 AAM78421	Aam78421 Human pro
25	4239.5	61.2	1306	4 AAB68218	Aab68218 Amino aci

26	4195.5	60.5	1297	2 AAY55932	Aay55932 Human ZC2
27	4021.5	58.0	1239	2 AAY55931	Aay55931 Human ZC1
28	3977	57.4	1212	4 AAE04368	Aae04368 Human kin
29	3977	57.4	1212	7 ADE25753	Ade25753 Human pro
30	3960.5	57.2	1165	3 AAB43016	Aab43016 Human ORF
31	3951.5	57.0	1233	2 AAY55954	Aay55954 Mouse STE
32	3872.5	55.9	1165	4 ABG17470	Abg17470 Novel hum
33	3872.5	55.9	1165	5 AAO18509	AAO18509 Human ins
34	3872.5	55.9	1165	6 ASU03499	Asu03499 Angiogene
35	3824	55.2	1175	4 ABG17468	Abg17468 Novel hum
36	3824	55.2	1175	5 AAU76512	Aau76512 Human HPK
37	3821	55.1	792	7 ADE34152	Ade34152 Human mis
38	3807.5	55.0	1655	4 ABG17466	Abg17466 Novel hum
39	3777.5	54.5	1135	3 AAY68784	Aay68784 Amino aci
40	3694.5	53.3	1219	4 ABG17469	Abg17469 Novel hum
41	2708.5	39.1	1109	2 AAY55953	Aay55953 Nematode
42	2280	32.9	510	3 AAB54144	Aab54144 Human pan
43	1862.5	26.9	1227	2 AAY55965	Aay55965 Full leng
44	1819.5	26.3	1581	5 AAE24145	Aae24145 Human kin
45	1686	24.3	425	4 AAB92792	Aab92792 Human pro

ALIGNMENTS

RESULT 1

ADE34150
ID ADE34150 standard; protein; 1312 AA.

AC ADE34150;

DT 29-JAN-2004 (first entry)

DE Human misshapen/NIKS-related kinase, Mink3a.

KW Human: misshapen/NIKS-related kinase; Mink3a; enzyme; antiinflammatory;
KW immunosuppressive; cytostatic; germinal centre kinase;
KW c-JUN N-terminal kinase; JNK; extracellular signal response kinase; ERK;
KW growth factor induced-ERK activation; proliferation;
KW cell proliferation disorder; cell survival;
KW intracellular signal transduction; apoptosis; morphological change;
KW cell migration; gene therapy; inflammatory disease; autoimmune disease;
KW immunodeficiency; cancer.

XX Homo sapiens.

OS US2003077597-A1.

PN 24-APR-2003.

PD 19-OCT-2001; 2001US-00029115.

PR 19-OCT-2001; 2001US-00029115.

PA (LUOY/) LUO Y.

PA (FUCA/) FU C A.

PA (SHEN/) SHEN M.

PI Luo Y, Fu CA, Shen M;

DR WPI; 2003-635076/60.

DR N-PSDB; ADE34151.

XX New misshapen/NIKS-related kinase nucleic acids and proteins useful in gene therapy and for treating disorders, e.g. acute and chronic inflammatory diseases.

PS Claim 5; SEQ ID NO 2; 53pp; English.

XX The invention relates to a recombinant nucleic acid capable of

hybridising to a Human DNA encoding misshapen/NIKS-related kinase

(Mink3a, 3b and 3c, germinal centre kinase proteins) appearing as

CC ADE34151, ADE34153 and ADE34155, or at least 90% identity to them, or

CC their complements. Also included are a recombinant polypeptide at least
 CC 95 % sequence identity to Mink3a, 3b or 3c (appearing as ADE34150,
 CC ADE34152 and ADE34154), screening for a candidate bioactive agent capable
 CC of modulating c-JUN N-terminal kinase (JNK) or extracellular signal
 CC response kinase (ERK) phosphorylation or activity, screening for a
 CC candidate bioactive agent capable of modulating growth factor induced-ERK
 CC activation in a mammalian cell, screening for a candidate bioactive agent
 CC capable of modulating proliferation in a mammalian cell, diagnosing a
 CC mammalian cell proliferation disorder, a medicament for treating a
 CC mammalian cell proliferation disorder and screening for a candidate agent
 CC capable of modulating cell survival. The MINK3 (mishapen/NIKs-related
 CC kinase) nucleic acids are useful in the modulation of intracellular
 CC signal transduction, cell proliferation, apoptosis, morphological change
 CC and migration of mammalian cells. MINK3 nucleic acids and proteins are
 CC specifically useful in gene therapy, and for treating, preventing or
 CC diagnosing acute and chronic inflammatory diseases, autoimmune diseases
 CC and diseases characterised by immunodeficiency. The compositions may also
 CC be used to treat MINK3 dysfunction related disorders, e.g. cancer. The
 CC nucleotide sequences may also be used as hybridisation probes, in
 CC chromosome and gene mapping, and in generating antisense RNA and DNA. The
 CC present sequence represents Mink3a.

XX Sequence 1312 AA;

Query Match 100.0%; Score 6929; DB 7; Length 1312;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDPAARSDDIDLSALRDPAGIFELVEVVGNGTYGYQYKGRHVKTGQLAAIKVMDYTE 60
 DB 1 MGDPAARSDDIDLSALRDPAGIFELVEVVGNGTYGYQYKGRHVKTGQLAAIKVMDYTE 60
 QY 61 DEEBEIKQINMLKKYSHRNATYGYAFIKKSPGNDQDLWVMEFCGAGSVTDLVKNT 120
 DB 61 DEEBEIKQINMLKKYSHRNATYGYAFIKKSPGNDQDLWVMEFCGAGSVTDLVKNT 120
 QY 121 KGNALKEDCIAYICREILGLAHLHAHKVIRHDIKQNVLLTENAEEVKLVDFGVAQLDR 180
 DB 121 KGNALKEDCIAYICREILGLAHLHAHKVIRHDIKQNVLLTENAEEVKLVDFGVAQLDR 180
 QY 181 TVGRNTFTGTYYWAVEVIACDENPDATYDYSDIWSLIGITAIEMAEAGAPLCDMPMR 240
 DB 181 TVGRNTFTGTYYWAVEVIACDENPDATYDYSDIWSLIGITAIEMAEAGAPLCDMPMR 240
 QY 241 ALFLIPRNPPLKSKKSKKIDIDTCLIKTYLSRPTEOLLKFPFIRDOPTEROVRI 300
 DB 241 ALFLIPRNPPLKSKKSKKIDIDTCLIKTYLSRPTEOLLKFPFIRDOPTEROVRI 300
 QY 301 QLKDHIDRSKRGKEETEYYSGEEDDSHGEGEPSSIMNVPGESTLRRREFLRLLQ 360
 DB 301 QLKDHIDRSKRGKEETEYYSGEEDDSHGEGEPSSIMNVPGESTLRRREFLRLLQ 360
 QY 361 ENKNSSEALKKQQQQLKQQQORPEAHIKHLLHQRRRIIEQKEERRRVEEQORRRERQK 420
 DB 361 ENKNSSEALKKQQQQLKQQQORPEAHIKHLLHQRRRIIEQKEERRRVEEQORRRERQK 420
 QY 421 LOEKQORLEDMQALRREERBRQAEQYKRLQLEQORQSERLQROLQQSHAYLKSILQ 480
 DB 421 LOEKQORLEDMQALRREERBRQAEQYKRLQLEQORQSERLQROLQQSHAYLKSILQ 480
 QY 481 QQQQQQQQLKQQQQLKPGDKRPLVHYGRGMNPAKPAWAREVEERTMKNQONSPLAKS 540
 DB 481 QQQQQQQQLKQQQQLKPGDKRPLVHYGRGMNPAKPAWAREVEERTMKNQONSPLAKS 540
 QY 541 KPGSTGPEPPIQASGPPGPIQSOTPMQRPVEPQEGPHKSLQDQPTNLNLAAPASHDPD 600
 DB 541 KPGSTGPEPPIQASGPPGPIQSOTPMQRPVEPQEGPHKSLQDQPTNLNLAAPASHDPD 600
 QY 601 PAIPAPTATPSARGAVIRQNSDPTSEGGPGSPNPFAWRPDNEAPPKVPQRTSSATATLN 660
 DB 601 PAIPAPTATPSARGAVIRQNSDPTSEGGPGSPNPFAWRPDNEAPPKVPQRTSSATATLN 660
 QY 661 TSGAGGSRPAQAVRARPRNSAWQIYLQRAERGTPKPPGPPAQPQPPNASSNPDLRRS 720

DB 661 TSGAGGSRPAQAVRARPRNSAWQIYLQRAERGTPKPPGPPAQPQPPNASSNPDLRRS 720
 QY 721 DPGHERSDVLPASHGHLPOAGSLERNRVGASSKLDSSPVLSPGNKAKPDDHRSRPGRPA 780
 DB 721 DPGHERSDVLPASHGHLPOAGSLERNRVGASSKLDSSPVLSPGNKAKPDDHRSRPGRPA 780
 QY 781 DFVLLKERTLEADPRPPKAMDYSSSSSEVESSDEDEEGEGGPAEGSRDTPGGRSDGDT 840
 DB 781 DFVLLKERTLEADPRPPKAMDYSSSSSEVESSDEDEEGEGGPAEGSRDTPGGRSDGDT 840
 QY 841 DSVSTMVVDHVEEITGTQPPYGGGTVMVQVQRTPEERNLLHADSNGYTNLPDVQPSHSPT 900
 DB 841 DSVSTMVVDHVEEITGTQPPYGGGTVMVQVQRTPEERNLLHADSNGYTNLPDVQPSHSPT 900
 QY 901 ENSKGQSPPSKDGSDGYSRGLVKAPGKSSFTMFVDLGIYQPGSGSDSIPITALVGGEGT 960
 DB 901 ENSKGQSPPSKDGSDGYSRGLVKAPGKSSFTMFVDLGIYQPGSGSDSIPITALVGGEGT 960
 QY 961 RLDQLQYDVRKGSVVNVNPTNTRAHSETPEIRKYKRFNSILCAALMGVNLVGTENGL 1020
 DB 961 RLDQLQYDVRKGSVVNVNPTNTRAHSETPEIRKYKRFNSILCAALMGVNLVGTENGL 1020
 QY 1021 MLLDRSGGKYVGLIGRRRRFQOMDVLBNLLITISGRNKLRYVYLSWLRNKLHNDPE 1080
 DB 1021 MLLDRSGGKYVGLIGRRRRFQOMDVLBNLLITISGRNKLRYVYLSWLRNKLHNDPE 1080
 QY 1081 VEKKQGTWTVGDMEGCGHYRVVKYERIKFLVIALKSSVEVYAWAPKPYHKFMAPKSFADL 1140
 DB 1081 VEKKQGTWTVGDMEGCGHYRVVKYERIKFLVIALKSSVEVYAWAPKPYHKFMAPKSFADL 1140
 QY 1141 PHRPLLDVLTVEEGORLKVYSSAGFAHVDVDSGNSYDIYIPVHIQSQITPHAIIFLPN 1200
 DB 1141 PHRPLLDVLTVEEGORLKVYSSAGFAHVDVDSGNSYDIYIPVHIQSQITPHAIIFLPN 1200
 QY 1201 TDGMELLCYDEDEGVYNTYGRRIIKDVVLQWGMPTSVAYICSNQIMGWGEKAIEIRSV 1260
 DB 1201 TDGMELLCYDEDEGVYNTYGRRIIKDVVLQWGMPTSVAYICSNQIMGWGEKAIEIRSV 1260
 QY 1261 TGHLDGVFMHKAQRLKFLCERNDKVFFASVRSVSGSSQVYFMTLNRNCIMNW 1312
 DB 1261 TGHLDGVFMHKAQRLKFLCERNDKVFFASVRSVSGSSQVYFMTLNRNCIMNW 1312
 RESULT 2
 AAO18508
 ID AAO18508 standard; protein; 1295 AA.
 AC AAO18508;
 XX
 DT 11-OCT-2002 (first entry)
 XX
 DE Human insulin receptor signaling modifier SEQ ID NO: 38.
 XX
 KW Human; insulin receptor signaling; insulin receptor signaling modifier;
 KW ISM; diabetes; metabolic syndrome; antidiabetic.
 XX
 OS Homo sapiens.
 XX
 FN WO200255664-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 11-JAN-2002; 2002WO-US001048.
 XX
 PR 12-JAN-2001; 2001US-0261226P.
 PR 12-JAN-2001; 2001US-0261303P.
 PR 12-JAN-2001; 2001US-0261304P.
 PR 12-JAN-2001; 2001US-0261335P.
 PR 12-JAN-2001; 2001US-0261336P.
 PR 12-JAN-2001; 2001US-0261361P.
 PR 12-JAN-2001; 2001US-0261456P.
 PR 12-JAN-2001; 2001US-0261457P.

PR 12-JAN-2001; 2001US-0261458P.
PR 12-JAN-2001; 2001US-0261459P.
PR 12-JAN-2001; 2001US-0261461P.
PR 12-JAN-2001; 2001US-0261518P.
PR 12-JAN-2001; 2001US-0261531P.
PR 12-JAN-2001; 2001US-0261532P.
PR 12-JAN-2001; 2001US-0261589P.
PR 12-JAN-2001; 2001US-0261590P.
PR 12-JAN-2001; 2001US-0261694P.
PR 12-JAN-2001; 2001US-0261695P.
PR 12-JAN-2001; 2001US-0261697P.
XX (EXEL-) EXELIXIS INC.
XX Seidel-Dugan C, Ferguson KC, Kidd T;
XX N-PSDB; AAL48627.
XX WPI; 2002-599664/64.
XX DR N-PSDB; AAL48627.
XX PT Identifying an insulin receptor signaling modulator, useful as drug
XX targets for treating diabetes or metabolic disorders, comprises
XX contacting an assay system comprising insulin receptor signaling
XX modifiers with a test agent.
XX Disclosure; Page 120-125; 232pp; English.
XX The present invention relates to a method of identifying a candidate
XX insulin receptor (INR) signaling modulating agent, involving contacting
XX an assay system comprising an insulin receptor signaling modifier (ISM)
XX polypeptide or nucleic acid with a test agent, and detecting a test agent
XX biased activity of the assay system. The method is useful for
XX identifying candidate INR signaling modulating agents. ISM genes may be
XX used as drug targets for treatment of disorders related to INR signaling
XX such as diabetes or metabolic syndrome. ISM nucleic acids and
XX polypeptides are useful for identifying and testing agents that modulate
XX ISM function and for other applications related to the involvement of ISM
XX in INR signaling, and for identifying subjects having a predisposition to
XX such diseases associated with INR signaling. The present sequence is an
XX ISM protein described in the exemplification of the invention
XX SQ Sequence 1295 AA;
Query Match 96.1%; Score 6660.5; DB 5; Length 1295;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 1273; Conservative 0; Mismatches 2; Indels 57; Gaps 2;
QY 1 MGDPAKSLDDIDLSALRDPAGIFELVEVNGTYGVYKGRHVKTGQLAAIKVMDYTE 60
DB 1 MGDPAKSLDDIDLSALRDPAGIFELVEVNGTYGVYKGRHVKTGQLAAIKVMDYTE 60
QY 61 DEEBIKQEIINMLKKYSHRNATYTGAFIKKSPGNDQDLWLVMEFCAGSVTLVKNT 120
DB 61 DEEBIKQEIINMLKKYSHRNATYTGAFIKKSPGNDQDLWLVMEFCAGSVTLVKNT 120
QY 121 KGNALKEDCIAYICREILRGLAHLHAHVHRIKQNVLLTENAELVKLVDFGVSQAQLDR 180
DB 121 KGNALKEDCIAYICREILRGLAHLHAHVHRIKQNVLLTENAELVKLVDFGVSQAQLDR 180
QY 161 TVGRNTFTGTPYMAPEVACDENPDATYDYSIDWSIGITAIEMASGAPPLCDMHPMR 240
DB 161 TVGRNTFTGTPYMAPEVACDENPDATYDYSIDWSIGITAIEMASGAPPLCDMHPMR 240
QY 241 ALFLIPRNPPLKSKGSKKFIIDFIDCLIKTYLSRPPTEQLLKFPIRQDPTQVRIR 300
DB 241 ALFLIPRNPPLKSKGSKKFIIDFIDCLIKTYLSRPPTEQLLKFPIRQDPTQVRIR 300
QY 301 QUKOHIDRSRKRKGKEETEYISGSEEDDGHGEGEPSSIMNVPGESTLRREFLRQQ 360
DB 301 QUKOHIDRSRKRKGKEETEYISGSEEDDGHGEGEPSSIMNVPGESTLRREFLRQQ 360
QY 361 ENKNSSEALKKQQQKQQQOORRPEAHIKHLLHQRORRREEOKEERRRVEEQORREREOK 420
DB 361 ENKNSSEALKKQQQKQQQOORRPEAHIKHLLHQRORRREEOKEERRRVEEQORREREOK 420

QY 421 LOEKEQORRLDMOALRRERROAEREOEYKRYKQLEBEOQSERLQORLOQEHAYLKSLO 480
DB 421 LOEKEQORRLDMOALRRERROAEREOEYKRYKQLEBEOQSERLQORLOQEHAYLKSLO 480
QY 481 QQQQQQQLOKQQQQQQLLPGDRKPLYHYGRGNWNPADKPAWAREVEERTMKNQKNSPLAKS 540
DB 481 QQQQQQQLOKQQQQQQLLPGDRKPLYHYGRGNWNPADKPAWAREVEERTMKNQKNSPLAKS 540
QY 541 KPGSTGPEPPIPOASPGPPGPLSQTPPMQRPVEPQEGPHK----- 580
DB 541 KPGSTGPEPPIPOASPGPPGPLSQTPPMQRPVEPQEGPHKSLVAHRVPLKPYAAPVPRSQ 600
QY 581 SLQDQPTRNLAAFPASHDPPDAIPAPTATPSARGAVIRONSDDPTSEGPPSPNPAAWRP 640
DB 601 SLQDQPTRNLAAFPASHDPPDAIPAPTATPSARGAVIRONSDDPTSEGPPSPNPAAWRP 660
QY 641 DNEAPPKVPQORTSSIALNTSGAGSRPAQAVRPRNSAWQIYLQRRARERTPKPPG 700
DB 661 DNEAPPKVPQORTSSIALNTSGAGSRPAQAVR----- 695
QY 701 PPAQPPGPPNASSNPDLRRSDPGWERSDVLPAASHGLPQAGSLERNRVGASSKLDSPV 760
DB 696 -----SNFDLRRSDPGWERSDVLPAASHGLPQAGSLERNRVGVSSKPDSPV 743
QY 761 LSPGNKAKPDDHRSRPGRPADFVLLKERTLDEAPPPKAMDYSSSESEVEDDEEG 820
DB 744 LSPGNKAKPDDHRSRPGRPADFVLLKERTLDEAPPPKAMDYSSSESEVEDDEEG 803
QY 821 EGGPAEGRSDTPGGRSDGTDVSSTVMVHDVEETGTQPPYGGTMMVQRTPEERNLLH 880
DB 804 EGGPAEGRSDTPGGRSDGTDVSSTVMVHDVEETGTQPPYGGTMMVQRTPEERNLLH 863
QY 881 ADSNGYTNLPDVQPSHSPSTENSKGQSPPSKQSGDYQSRGLVKAPGKSSFTMFVLDGIY 940
DB 864 ADSNGYTNLPDVQPSHSPSTENSKGQSPPSKQSGDYQSRGLVKAPGKSSFTMFVLDGIY 923
QY 941 QPGSGSDSIPITLVGGEGTRLDQLDYDVRKGSVNVNPTNTRAHSETPEIRKYKRRNS 1000
DB 924 QPGSGSDSIPITLVGGEGTRLDQLDYDVRKGSVNVNPTNTRAHSETPEIRKYKRRNS 983
QY 1001 EILCAALWGNVLVGTENGLMLDLSRSGKGVYGLIGRRRFOQMDVLEGLNLLITISGRN 1060
DB 984 EILCAALWGNVLVGTENGLMLDLSRSGKGVYGLIGRRRFOQMDVLEGLNLLITISGRN 1043
QY 1061 KLRVYLSWLRNKLHNDPEVEKKQGTTVGDMEGCGHYRVVYKRIKFLVIALKSSVEV 1120
DB 1044 KLRVYLSWLRNKLHNDPEVEKKQGTTVGDMEGCGHYRVVYKRIKFLVIALKSSVEV 1103
QY 1121 YAWAPKPYHKPMAPKSPADLPHRPLLDLTVBEGQRLKVIYSSAGHADVDSGNSYDI 1180
DB 1104 YAWAPKPYHKPMAPKSPADLPHRPLLDLTVBEGQRLKVIYSSAGHADVDSGNSYDI 1163
QY 1181 YIPVHIQSQTTPHAIIFLPTNDGMELLCYDEGEVYVNTYGRRIKDVVLQWGMPTSVAY 1240
DB 1164 YIPVHIQSQTTPHAIIFLPTNDGMELLCYDEGEVYVNTYGRRIKDVVLQWGMPTSVAY 1223
QY 1241 ICSNQIMGWGKALFIRSVETGHLDDGVFMHKAQRLKFLCERNDKVFFASVRSGSSQVY 1300
DB 1224 ICSNQIMGWGKALFIRSVETGHLDDGVFMHKAQRLKFLCERNDKVFFASVRSGSSQVY 1283
QY 1301 FMTLNRNCIMNW 1312
DB 1284 FMTLNRNCIMNW 1295

RESULT 3
AAY55933
ID AAY55933 standard; protein; 1326 AA.
XX
AC AAY55933;
XX
DT 18-FEB-2000 (first entry)

Qy	1187	QSOITPHAIIFLPNTDGMEMLLCYEDSGVYNTYGRIIKDVVLQWGMPTSVAYICSNQI	1246
Db	1201	QSOITPHAIIFLPNTDGMEMLLCYEDSGVYNTYGRIIKDVVLQWGMPTSVAYICSNQI	1260
Qy	1247	MGWGEKAIEIRSVETGHLDGVMFHKRAQLKFLCERNDKVPFASVRSGGSSQYVFMTLNR	1306
Db	1261	MGWGEKAIEIRSVETGHLDGVMFHKRAQLKFLCERNDKVPFASVRSGGSSQYVFMTLNR	1320
Qy	1307	NCIMNW 1312	
Db	1321	NRINNW 1326	
RESULT 4			
AAM79153			
XX	AA	AAM79153 standard; protein; 1303 AA.	
XX	AC	AAM79153;	
XX	DT	06-NOV-2001 (first entry)	
XX	DE	Human protein SEQ ID NO 1815.	
XX	KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis;	
XX	KW	tissue growth factor; immunomodulatory; cancer; leukaemia;	
XX	KW	nervous system disorder; arthritis; inflammation.	
XX	OS	Homo sapiens.	
XX	PN	WO200157190-A2.	
XX	PD	09-AUG-2001.	
XX	PF	05-FEB-2001; 2001WO-US0004098.	
XX	PR	03-FEB-2000; 2000US-00496914.	
XX	PR	27-APR-2000; 2000US-00560875.	
XX	PR	20-JUN-2000; 2000US-00598075.	
XX	PR	19-JUL-2000; 2000US-00620325.	
XX	PR	01-SEP-2000; 2000US-00654936.	
XX	PR	15-SEP-2000; 2000US-00663561.	
XX	PR	20-OCT-2000; 2000US-00693325.	
XX	PR	30-NOV-2000; 2000US-00728422.	
XX	PA	(HYSE-) HYSEQ INC.	
XX	PI	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;	
XX	PI	Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;	
XX	PI	Xue Aa, Yang Y, Wejhrman T, Goodrich R;	
XX	DR	WPI; 2001-476283/51.	
XX	DR	N-PSDB; AAKS2286.	
XX	PT	Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy.	
XX	PS	Claim 20; Page 4185-4188; 6221pp; English.	
XX	CC	The invention relates to polynucleotides (AAKS51456-AAKS3435) and the encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AAKS2581), 2111 (AAKS2582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication	

QY 121 KGNALKEDCIAYICREILRGLAHLHAHVKVIRHDIKQNVLLTENAEEVKLVDFGVAQQLDR 180
DB 121 KGNALKEDCIAYICREILRGLAHLHAHVKVIRHDIKQNVLLTENAEEVKLVDFGVAQQLDR 180
QY 181 TVGRENFTIGTYPWMAPEVACDENPDATYDVRSDIWSLIGITAIEMASGAPPLCDMHPWR 240
DB 181 TVGRENFTIGTYPWMAPEVACDENPDATYDVRSDIWSLIGITAIEMASGAPPLCDMHPWR 240
QY 241 ALFLIPRNPPLRLSKSKWKKFIDFIDTCLIKTYLSRPPTEQLLKFPPIRDOPTERQVRI 300
DB 241 ALFLIPRNPPLRLSKSKWKKFIDFIDTCLIKTYLSRPPTEQLLKFPPIRDOPTERQVRI 300
QY 301 QLKOHIDRSRKRGKEETEYYSGBEDDSSHGEGPSSIMNVPGESTLRREPLRLQQ 360
DB 301 QLKOHIDRSRKRGKEETEYYSGBEDDSSHGEGPSSIMNVPGESTLRREPLRLQQ 360
QY 361 ENKSNSEALKKQOQQLQOQOQDORPEAHIKHLHQRORRIEEOKEERRRVEEOORREROR 420
DB 361 ENKSNSEALKKQOQQLQOQOQDORPEAHIKHLHQRORRIEEOKEERRRVEEOORREROR 420
QY 421 LOEKEQORRLEDMQALREERERQAREQ-----EY 451
DB 421 LOEKEQORRLEDMQALREERERQAREQEYIRHLEBEQROLETLQOQLLEQALLLEY 480
QY 452 KRKQLEEQORSERLQOQEHAYILKSLQOQOQOQOQOQOQOQLLPGDRKPLHYHGRGM 511
DB 481 KRKQLEEQORSERLQOQEHAYILKSLQOQOQOQOQOQOQOQLLPGDRKPLHYHGRGM 540
QY 512 NPADKPAAWAVEERTRNWKOQNSPLAKSKGCTGPEPPIPOASPGPGPLSOTPPMORP 571
DB 541 NPADKPAAWAVEERTRNWKOQNSPLAKSKGCTGPEPPIPOASPGPGPLSOTPPMORP 600
QY 572 VEQSGPHK-----SLQOQTRNLAAFPASHDDPDAIPATATPS 611
DB 601 VEQSGPHKSLVAHVPLKPYAAPVPRSQSLQOQTRNLAAFPASHDDPDAIPATATPS 660
QY 612 ARGAVIRQNSDPTSEGGPSPNPAAWVRPDNEAPPKVPFORTSSITATLNTSGAGSRPAQ 671
DB 661 ARGAVIRQNSDPTSEGGPSPNPAAWVRPDNEAPPKVPFORTSSITATLNTSGAGSRPAQ 720
QY 672 AVRARPNSAWQIYLQRAERGTGPKPGPPAQPFGPPNASSNDLRRSDPGWERSDVL 731
DB 721 AVRA-----SNPDLRRSDPGWERSDVL 743
QY 732 PASHGHLPOAGSLERNRVGASSKLDSPVLSGKAKKPDHRSRGRPA-----DFV 783
DB 744 PASHGHLPOAGSLERNRVGVSSKPDSSPVLSPGNKAKPDHRSRGRPASVKRAIGEDFV 803
QY 784 LKERTLDEAPPPKAMDYSSSEBESSDEDEEGEGGPAEGSRDTPGGRSDGDTDSV 843
DB 804 LKERTLDEAPPPKAMDYSSSEBESSDEDEEGEGGPAEGSRDTPGGRSDGDTDSV 863
QY 844 STMVHVDVEEITGTOPPGGTMVQVTPPEERNLLHADSNGYTNLPDVQVPSHSPTENS 903
DB 864 STMVHVDVEEITGTOPPGGTMVQVTPPEERNLLHADSNGYTNLPDVQVPSHSPTENS 923
QY 904 KQSGPPSKDGSQDYSRGLVAPGKSSFTMFVDLGIYQPGSGDSIPITLVGGSGTRLD 963
DB 924 KQSGPPSKDGSQDYSRGLVAPGKSSFTMFVDLGIYQPGSGDSIPITLVGGSGTRLD 983
QY 964 QLYQDVVRKGSVVNVNFTNTRAHSETPEIRKYKRRFNSBILCAALMGVNLVGTENGLMLL 1023
DB 984 QLYQDVVRKGSVVNVNFTNTRAHSETPEIRKYKRRFNSBILCAALMGVNLVGTENGLMLL 1043
QY 1024 DRSGGQKVGILGRRRFOQMDVLEGNLLITISGRNKLVRVYLSWLRNKLHNDPEVEK 1083
DB 1044 DRSGGQKVGILGRRRFOQMDVLEGNLLITISGRNKLVRVYLSWLRNKLHNDPEVEK 1103
QY 1084 KQGWTTVGDMECGHYRVVKYERIKFLVLTALKSSVEVWAPKPKYHKFWAFKSPADLPHR 1143
DB 1104 KQGWTTVGDMECGHYRVVKYERIKFLVLTALKSSVEVWAPKPKYHKFWAFKSPADLPHR 1163

QY 1144 PLLVDLTVESGQRLKVIYGSAGFAHVDVDSGNSYDIYIPVHIQSQITPHAIIFLPNTD 1203
DB 1164 PLLVDLTVESGQRLKVIYGSAGFAHVDVDSGNSYDIYIPVHIQSQITPHAIIFLPNTD 1223
QY 1204 MEMLLCYEDEGVVYNTYGRIRIKOVVLQWGMPTSVAVYCSNQIMGWGEKAIEIRSVETGH 1263
DB 1224 MEMLLCYEDEGVVYNTYGRIRIKOVVLQWGMPTSVAVYCSNQIMGWGEKAIEIRSVETGH 1283
QY 1264 LDGVFMHKAQRLKFLCERNDKVFFASVRSQSSQVYFMTLNRNCIMNW 1312
DB 1284 LDGVFMHKAQRLKFLCERNDKVFFASVRSQSSQVYFMTLNRNCIMNW 1332
RESULT 8
ADE34154
ID ADE34154 standard; protein; 1276 AA.
XX
AC ADE34154;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human misshapen/NIKs-related kinase, Mink3c.
XX
KW Human; misshapen/NIKs-related kinase; Mink3c; enzyme; antiinflammatory;
KW immunosuppressive; cytostatic; germinal centre kinase;
KW c-JUN N-terminal kinase; JNK; extracellular signal response kinase; ERK;
KW growth factor induced-ERK activation; proliferation;
KW cell proliferation disorder; cell survival;
KW intracellular signal transduction; apoptosis; morphological change;
KW cell migration; gene therapy; inflammatory disease; autoimmune disease;
KW immunodeficiency; cancer.
XX
OS Homo sapiens.
XX
PN US2003077597-A1.
XX
PD 24-APR-2003.
XX
PF 19-OCT-2001; 2001US-00029115.
XX
PR 19-OCT-2001; 2001US-00029115.
XX
PA (LUOY/) LUO Y.
PA (FUCA/) FU C A.
PA (SHEN/) SHEN M.
XX
PI Luo Y, Fu CA, Shen M;
XX
PI WPI; 2003-635076/60.
DR N-PSDB; ADE34155.
XX
PT New misshapen/NIKs-related kinase nucleic acids and proteins useful in
PT gene therapy and for treating disorders, e.g. acute and chronic
PT inflammatory diseases.
XX
PS Claim 5; SEQ ID NO 6; 53pp; English.
XX
CC The invention relates to a recombinant nucleic acid capable of
CC hybridising to a Human DNA encoding misshapen/NIKs-related kinase
CC (Mink3a, 3b and 3c, germinal centre kinase proteins) appearing as
CC ADE34151, ADE34153 and ADE34155, or at least 90% identity to them, or
CC their complements. Also included are a recombinant polypeptide at least
CC 95 % sequence identity to Mink3a, 3b or 3c (appearing as ADE34150,
CC ADE34152 and ADE34154), screening for a candidate bioactive agent capable
CC of modulating c-JUN N-terminal kinase (JNK) or extracellular signal
CC response kinase (ERK) phosphorylation or activity, screening for a
CC candidate bioactive agent capable of modulating growth factor induced-ERK
CC activation in a mammalian cell, screening for a candidate bioactive agent
CC capable of modulating proliferation in a mammalian cell, diagnosing a
CC mammalian cell proliferation disorder, a medicament for treating a
CC mammalian cell proliferation disorder and screening for a candidate agent
CC capable of modulating cell survival. The MINK3 (misshapen/NIKs-related
CC kinase) nucleic acids are useful in the modulation of intracellular

PR 25-FEB-2000; 2000US-0184951P.
PR 28-FEB-2000; 2000US-0185548P.
PR 01-MAR-2000; 2000US-0185967P.
PR 18-APR-2000; 2000US-0197723P.
PR 27-APR-2000; 2000US-0199957P.
PR 23-FEB-2001; 2001US-00789390.
XX
XX (CURA-) CURAGEN CORP.
PI Vernet CAM, Fernandes E, Shinkets RA, Macdougall J, Spaderna SK;
XX WPI; 2001-582051/65.
XX N-PSDB; AAD17761.
XX
XX New isolated KIAA1233-like, STE20-like, or trypsin inhibitor-like
PT polypeptide for diagnosing and treating pathological disorders, such as
PT Parkinson's disease and for use in pharmacogenomics.
XX
XX Claim 1; Page 59; 189pp; English.
XX
CC The invention relates to novel human polypeptides referred as NOV-X and
CC their corresponding nucleic acid sequences. NOV-X collectively include
CC NOV-1, NOV-2a and NOV-2b which are novel KIAA1233-like polypeptides, NOV-
CC 3a, NOV-3b, NOV-3c and NOV-3d which are novel STE20-like polypeptides and
CC NOV-4a, NOV-4b, NOV-4c, NOV-4d and NOV-4e which are novel trypsin
CC inhibitor-like polypeptides. NOV-X is used to identify a potential
CC therapeutic agent that can modulate its activity and can be used for
CC treating a pathology related to aberrant expression or aberrant
CC physiological interactions of NOV-X. NOV-X or its DNA is used to
CC determine the presence or predisposition to a disease associated with
CC altered levels of NOV-X. NOV-X, its DNA and its antibody are used to
CC treat or prevent a pathology associated with NOV-X. The pathological
CC states that can be treated or prevented are haematopoietic, cancer,
CC immunological, tumour, neurodegenerative (e.g. Alzheimer's and
CC Parkinson's disease), human immunodeficiency virus (HIV) illness and
CC fertility disorders. NOV-X and its DNA are used in pharmacogenomics for
CC predictive medicine. NOV-X DNA is used in gene therapy. The present
CC sequence is human novel STE20-like protein, NOV-3c. NOV-3c gene is
CC located at chromosome 17
XX
XX Sequence 1273 AA;
SQ
Query Match 91.0%; Score 6304.5; DB 4; Length 1273;
Best Local Similarity 90.5%; Pred. No. 0;
Matches 1221; Conservative 2; Mismatches 13; Indels 113; Gaps 5;
QY 1 MGDPAFSLDDIDLSALRDPAGIFELVEVWGNGTYGVYKGRHVKTGQLAAIKVMDYTE 60
DB 1 MGDPAFSLDDIDLSALRDPAGIFELVEVWGNGTYGVYKGRHVKTGQLAAIKVMDYTE 60
QY 61 DEEEIKQEIINMLKKYSHRNATYTGAFIKKSPGNDQDLWLVMEFCAGSVTLVKNT 120
DB 61 DEEEIKQEIINMLKKYSHRNATYTGAFIKKSPGNDQDLWLVMEFCAGSVTLVKNT 120
QY 121 KGNALKEDCIAYCREILRGLAHLAHKVIHRDIKQNVLLTENAELVKLVFGVSAQLDR 180
DB 121 KGNALKEDCIAYCREILRGLAHLAHKVIHRDIKQNVLLTENAELVKLVFGVSAQLDR 180
QY 181 TVGRNTFTGTYYMAPEVIACDENPDATYDRSDIWSLIGITAIEMAGAPLCDHMPR 240
DB 181 TVGRNTFTGTYYMAPEVIACDENPDATYDRSDIWSLIGITAIEMAGAPLCDHMPR 240
QY 241 ALFLIPRPPPLKSKWKKFIDTICLIKTYLSRPPTEQLLKFPPIRDIQPTROVRI 300
DB 241 ALFLIPRPPPLKSKWKKFIDTICLIKTYLSRPPTEQLLKFPPIRDIQPTROVRI 300
QY 301 QLKDHIDRSRKRGKEETFEYSGSEEDDSHGEGEPSSIMNVPGSESTLRREFLRQQ 360
DB 301 QLKDHIDRSRKRGKEETFEYSGSEEDDSHGEGEPSSIMNVPGSESTLRREFLRQQ 360
QY 361 ENKNSSEALKKQOOQLQOOOQDPEAHIKHLHQRORRIEEOKEERRRVEEQORREORK 420
DB 361 ENKNSSEALKKQOOQLQOOOQDPEAHIKHLHQRORRIEEOKEERRRVEEQORREORK 420

QY 421 LOEKEQORRLDQALRRERROAEREQ-----EY 451
DB 421 LOEKEQORRLDQALRRERROAEREQEYIRHRLBEEQOLETLOOQLQEQALLLEY 480
QY 452 KRKOLEQORSERLQRLQOEHAHAYLSLQOOOQOQLOKQOOOQLLPDRKPLVHYGRGM 511
DB 481 KRKOLEQORSERLQRLQOEHAHAYLSLQOOOQOQLOKQOOOQLLPDRKPLVHYGRGM 540
QY 512 NPADKPAREVEERTMKNQONSPLAKSKGSGTGPPEPIPAQSPGPGPILSQTPPMQRP 571
DB 541 NPADKPAREVVAHR-----VPLKPYAAPVPRS----- 569
QY 572 VEOGEGHKSLODQOTRNLAAPASHDPDPAIPAPTATPSARGAVIRONSPTSEGGPGS 631
DB 570 -----QSLQDQOTRNLAAPASHDPDPAIPAPTATPSARGAVIRONSPTSEGGPGS 621
QY 632 PNPPAWVRPDNEAPKVPQRTSSITATLNTSGAGSRPAQAVRARPRNSAWQIYLQRR 691
DB 622 PNPPAWVRPDNEAPKVPQRTSSITATLNTSGAGSRPAQAVRA----- 665
QY 692 ERGTPKPPGPAQPPGPNASSNPDLRRSDPGWERSDVLPSASHGHLPAQGLERNRGA 751
DB 666 -----SNPDLRRSDPGWERSDVLPSASHGHLPAQGLERNRGA 704
QY 752 SSKLDSSPVLSPGNKAKPDDHRSRPPRPA-----DFVLLKERTLDEAPRPPKAMDY 803
DB 705 SSKLDSSPVLSPGNKAKPDDHRSRPPRPA-----DFVLLKERTLDEAPRPPKAMDY 764
QY 804 SSSSEEVESSEDDDEEGEGGPAEGSRDTPGGRSDGTDSTVSTVMVHDVEEITGTQPPYGG 863
DB 765 SSSSEEVESSEDDDEEGEGGPAEGSRDTPGGRSDGTDSTVSTVMVHDVEEITGTQPPYGG 824
QY 864 GTMVVQRTPEERNLLHADSNYTNLDPVQPSHPSTENSKGQSPSPKSGDQVSRGLV 923
DB 825 GTMVVQRTPEERNLLHADSNYTNLDPVQPSHPSTENSKGQSPSPKSGDQVSRGLV 884
QY 924 KAPGKSSFTMFVDLGIYQPGSGSDSIPTALVGGEGTGLDQLOQYDVRKGSVNVNPTNTR 983
DB 885 KAPGKSSFTMFVDLGIYQPGSGSDSIPTALVGGEGTGLDQLOQYDVRKGSVNVNPTNTR 944
QY 984 AHSETPEIRKYKGFNSEILCAALWGVNLLVGTENGMLLDRSGQKGYGLIGRRRFOOM 1043
DB 945 AHSETPEIRKYKGFNSEILCAALWGVNLLVGTENGMLLDRSGQKGYGLIGRRRFOOM 1004
QY 1044 DVLKGLNLLITISGRNKLRYVYLSWLNKILHNDPEVEKKQGWTTVGDMEGCGHYRVVK 1103
DB 1005 DVLKGLNLLITISGRNKLRYVYLSWLNKILHNDPEVEKKQGWTTVGDMEGCGHYRVVK 1064
QY 1104 YERIKFLVIALKSSVEVYAWAPKPYHKFMAFKSPADLPHRPLLVDLTVEEGQRLKVIYS 1163
DB 1065 YERIKFLVIALKSSVEVYAWAPKPYHKFMAFKSPADLPHRPLLVDLTVEEGQRLKVIYS 1124
QY 1164 SAGHAVDVDSGNSYDIYIPVHIQSOITPHAIIFLPNTDGMELLCYDEDEGVYNTYGR 1223
DB 1125 SAGHAVDVDSGNSYDIYIPVHIQSOITPHAIIFLPNTDGMELLCYDEDEGVYNTYGR 1184
QY 1224 IKDVVLWGEPMPTSVAVICNSQIMGWGEKATEIRSVETGHLDGVMHKAQRLKFLCBRN 1283
DB 1185 IKDVVLWGEPMPTSVAVICNSQIMGWGEKATEIRSVETGHLDGVMHKAQRLKFLCBRN 1244
QY 1284 DKVFFASVRSGSGSQVYFMTLNRNCIMNW 1312
DB 1245 DKVFFASVRSGSGSQVYFMTLNRNCIMNW 1273

RESULT 11
ABG24020
ID ABG24020 standard; protein; 1701 AA.
XX
AC ABG24020;
XX DT 18-FEB-2002 (first entry)


```

QY 1264 LDGVFMHKAQRLKFLCERNDKVFFASVRSKGSS 1297
DB 1166 LDGVFMHKAQRLKFLCERNDKVFFASVRSKGSS 1199

RESULT 12
ID AAB68221 standard; protein; 1324 AA.
XX AAB68221;
AC AAB68221;
XX 09-JUL-2001 (first entry)
XX Amino acid sequence of cell cycle protein Thnk isoform 6.
DE Cell cycle protein; Thnk; germinal center kinase; Nck; Traf2; cell cycle;
KW tumour necrosis factor receptor associated factor 2; gene therapy.
XX Homo sapiens.
XX OS
XX FH Key Location/Qualifiers
FT Misc-difference 1324
FT /note= "Glx encoded by TAA"
XX WO200129197-A2.
XX 26-APR-2001.
XX 23-OCT-2000; 2000WO-US041455.
XX 21-OCT-1999; 99US-00425324.
XX (RIGE-) RIGEL PHARM INC.
XX Luo Y, Fu CA, Shen M;
XX WPI; 2001-300333/31.
XX N-PSDB; AAF84944.
XX Novel germinal center kinase cell cycle polypeptide useful for screening
XX modulators of cell cycle and for use in diagnostics and therapeutics.
XX Claim 14; Fig 33; 96pp; English.
XX The present sequence represents an isoform of a human cell cycle protein
XX designated Thnk. Thnk is a member of the germinal center kinase family,
XX Thnk binds to tumour necrosis factor (TNF) receptor associated factor 2
XX (Traf2) or the adapter protein Nck. Thnk polypeptides and polynucleotides
XX are useful for screening for compounds that are modulators of cell cycle,
XX and in therapeutics. The Thnk polynucleotide is also useful in gene
XX therapy
XX SQ Sequence 1324 AA;

Query Match 63.1%; Score 4370.5; DB 4; Length 1324;
Best Local Similarity 64.7%; Pred. No. 9.7e-252;
Matches 899; Conservative 137; Mismatches 210; Indels 143; Gaps 31;

QY 1 MGDPAPARSLDDIDLSALRDPAGIFELVEVNGTYGQVYKGRHVKTGQLAAIKVMYTG 60
DB 1 MASDSPARSLDDIDLSALRDPAGIFELVEVNGTYGQVYKGRHVKTGQLAAIKVMYTG 60
QY 61 DEEEIKQBINLKKYSHRNATATYGAFFIKKSPGNDQDLWLVMEFCAGSVTDLVKNT 120
DB 61 DEEEIKQBINLKKYSHRNATATYGAFFIKKSPGNDQDLWLVMEFCAGSVTDLVKNT 120
QY 121 KGNALKEBCIAYICREILRGLAHLHAHKVYIHRDIKQGNVLLTENAEVKLVDFGVSAQLDR 180
DB 121 KGNLKEEWIAYICREILRGLSHLHQHKVYIHRDIKQGNVLLTENAEVKLVDFGVSAQLDR 180
QY 181 TVGRNTFTGTPYMAPEVIACDENPDATYDFKSDLSLGLTAIMAEGAPPLCDMHPMR 240
DB 181 TVGRNTFTGTPYMAPEVIACDENPDATYDFKSDLSLGLTAIMAEGAPPLCDMHPMR 240

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2004, 00:50:59 ; Search time 157 Seconds

(without alignments)
2636.687 Million cell updates/sec

Title: US-10-029-115-2

Perfect score: 6929

Sequence: 1 MDPAPARSLDDIDLSALRD.....SGGSSQVYFMTLNRNCINMW 1312

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertibrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6701	96.7	1334	11 Q7TT13	Q7tt13 mus musculus
2	4189.5	60.5	1303	13 Q7SY42	Q7sy42 brachydanio
3	2805.5	40.5	1550	5 Q9W002	Q9w002 drosophila
4	2698	38.9	1082	5 Q8T8M3	Q8t8m3 caenorhabdi
5	2696.5	38.9	1087	5 Q9XYC3	Q9xyc3 caenorhabdi
6	2692	38.9	1096	5 Q95Z17	Q95z17 caenorhabdi
7	2661	38.4	1072	5 Q814B5	Q814b5 caenorhabdi
8	1855	26.8	538	4 Q724L4	Q724l4 homo sapien
9	1826	26.4	1582	4 Q722Y5	Q722y5 homo sapien
10	1702.5	24.6	1455	11 Q9R0G8	Q9r0g8 mus musculus
11	1701.5	24.6	1455	11 Q9R0S4	Q9r0s4 mus musculus
12	1551	22.4	292	4 Q9HBM9	Q9hbm9 homo sapien
13	1458	21.0	792	5 Q9UAN7	Q9uan7 caenorhabdi
14	1377	19.9	728	5 Q9XYC4	Q9xyc4 caenorhabdi
15	993	14.3	334	11 Q9C9S9	Q9c9s9 mus musculus
16	829.5	12.0	169	11 Q61155	Q61155 mus musculus

17	811.5	11.7	1218	5	Q9V8R6	Q9v8r6 drosophila
18	804.5	11.6	1615	4	Q9NV98	Q9nv98 homo sapien
19	804.5	11.6	1616	4	Q8NEV4	Q8nev4 homo sapien
20	800.5	11.6	1613	11	Q8K3H5	Q8k3h5 mus musculus
21	795	11.5	1775	13	Q90XG6	Q90xg6 brachydanio
22	779	11.2	1838	13	Q9DG88	Q9dgg8 morone saxa
23	765	11.0	1310	13	Q800Q6	Q800q6 morone saxa
24	764.5	11.0	1192	4	Q8IX64	Q8ix64 homo sapien
25	764.5	11.0	1251	4	Q8IX68	Q8ix68 homo sapien
26	764.5	11.0	1275	4	Q8IX65	Q8ix65 homo sapien
27	764.5	11.0	1278	4	Q8WXR4	Q8wxr4 homo sapien
28	764.5	11.0	1314	4	Q8IX66	Q8ix66 homo sapien
29	764.5	11.0	1341	4	Q8IX67	Q8ix67 homo sapien
30	761.5	11.0	1113	4	Q96N94	Q96n94 homo sapien
31	734	10.6	737	5	Q86IX1	Q86ix1 dictyosteli
32	734	10.6	829	3	Q87LH9	Q87lh9 neurospora
33	720	10.4	825	10	Q8SAE1	Q8sae1 triticum mo
34	718	10.4	825	10	Q9ARL7	Q9arl7 hordeum vul
35	713.5	10.3	487	11	Q9J111	Q9j111 mus musculus
36	712	10.3	950	13	Q9VHC9	Q9vhc9 xenopus lae
37	711.5	10.3	842	10	Q9FNU3	Q9fnu3 oryza sativ
38	710	10.2	836	10	Q24527	Q24527 arabidopsis
39	707.5	10.2	1120	10	Q9LQA1	Q9lqa1 arabidopsis
40	703	10.1	491	13	Q802A6	Q802a6 squalus aca
41	701.5	10.1	497	11	Q9J110	Q9j110 mus musculus
42	701.5	10.1	539	11	Q80UG4	Q80ug4 mus musculus
43	701	10.1	947	5	Q8SYA1	Q8syal drosophila
44	700	10.1	651	5	Q95ZN6	Q95zn6 caenorhabdi
45	700	10.1	947	5	Q8MLI8	Q8mli8 drosophila

ALIGNMENTS

RESULT 1

Q7TT13 PRELIMINARY; PRT; 1334 AA.

AC Q7TT13; 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)

DE Hypothetical protein.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Frange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Faney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;

Db	886	QQSHSPGPTGLNSQDLSDMDEFLGG----	GSKASFTPPVDPVVQTSPTDDNDESSA	944
Qy	952	TALVGGEGRDLQIQV-DVRKGSVWVNPNTN	TRAHSETPEIRKYKGRFNSILCAALWGV	1011
Db	942	AMFANELLRQEQAKLNEARKISVWVNPNTN	IRPHSDTPEIRKYKGRFNSILCAALWGV	1003
Qy	1011	NLLVGTENGMLLDRSQGGKQVYGLIGRRRQ	QOMDVLGLNLLITISGKRNKLRVYYLSWL	1070
Db	1002	NLLVGTENGMLLDRSQGGKQVYLNLTIRRR	QOMDVLGLNVLVTISGKRNKLRVYYLSWL	1063
Qy	1071	RNKILHNDPEVEKKQGGTWTGDMGEGCHYR	VWKYERIKFLVIALKSSVEVYAWAPKPYHK	1133
Db	1062	RNRILHNDPEVEKKQGGTWTGELGCHYKVI	YERIKFLVIALKNSVEIYAWAPKPYHK	1122
Qy	1131	FMAFKSFADLPHRPLLVLTVEEGQRLKVIY	GSAGFHAVDVDSGNSYDVIYPVHIQSOI	1190
Db	1122	FMAFKSFTELQHRPQLVLTVEEGQRLKVIY	GSVGFHVLDVDSGSPYDIYPHSIQSTV	1187
Qy	1191	TPHAIIFLNTDGMEMLLCYDESGVYVNTY	GRIIKDVVLWGENMPTSVAYICSNQIMGW	1255
Db	1182	TPHAIIVLPKTDGMEMLLRYDESGVYVNTY	GRITKDVVLWGENMPTSVAYITHSNQTMGW	1244
Qy	1251	EKAIEIRSVETGHLDGVMFKRAQRLKFLC	RNDKVFASFVRSGSSQVYFMTLNRNCIM	1311
Db	1242	EKAIEIRSVETGHLDGVMFKRAQRLKFLC	RNDKVFASFVRSGSSQVFFMTLNRNSIM	1300
Qy	1311	NW 1312		
Db	1302	NW 1303		
RESULT 3				
Q9W002 PRELIMINARY; PRT; 1550 AA.				
ID	Q9W002	MEDLINE=20196006; PubMed=10731132;		
AC	Q9W002;	01-MAY-2000 (TREMBlrel. 13, Created)		
DT	01-OCT-2002	(TREMBlrel. 22, Last sequence update)		
DT	01-OCT-2003	(TREMBlrel. 25, Last annotation update)		
DE	CG16973 protein.			
GN	MSN OR CG16973.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
ON	NCBI_TaxID=7227;			
RX	[1]	SEQUENCE FROM N.A.		
RP	STRAIN=Berkeley;			
RC	ADAMS M.D., CELINKER S.E., HOLT R.A., EVANS C.A., GOCAYNE J.D.,			
RA	AMANTIDES P.G., SCHERER S.E., LI P.W., HOSKINS R.A., GALLE R.F.,			
RA	GEORGE R.A., LEWIS S.E., RICHARDS S., ASHBURNER M., HENDERSON S.N.,			
RA	SUTTON G.G., WORTMAN J.R., YANDELL M.D., ZHANG Q., CHEN L.X.,			
RA	BRANDON R.C., ROGERS Y.-H.C., BLAZEJ R.G., CHAMPE M., PREIFFER B.D.,			
RA	WAN K.H., DOYLE C., BAXTER E.G., HELT G., NELSON C.R., MIKLOS G.L.G.,			
RA	ABRIL J.F., AGBAYANI A., AN H.-J., ANDREWS-PFANNKUCH C., BALDWIN D.,			
RA	BALLOU R.M., BASU A., BAXENDALE J., BAYRAKTAROGU L., BEASLEY E.M.,			
RA	BEESON K.Y., BENOS P.V., BERMAN B.P., BHANDARI D., BOLSHAKOV S.,			
RA	BORKOVA D., BOTCHAN M.R., BOUCK J., BROKSTEIN P., BROTTIER P.,			
RA	BURTIS K.C., BUSAM D.A., BUTLER H., CADIEU E., CENTER A., CHANDRA I.,			
RA	CHERRY J.M., CAWLEY S., DAHLKE C., DAVENPORT L.B., DAVIES P.,			
RA	DE PABLOS B., DELCHER A., DENG Z., MAYS A.D., DEW I., DIETZ S.M.,			
RA	DODSON K., DOUP L.E., DOWNES M., DUGAN-ROCHA S., DUNKOV B.C., DUNN P.,			
RA	DURBIN K.J., EVANGELISTA C.C., FERRAZ C., FERRIERA S., FLEISCHMANN W.,			
RA	FOSLER C., GABRIELIAN A.E., GARG N.S., GELBART W.M., GLASSER K.,			
RA	GLODEK A., GONG F., GORRELL J.H., GU Z., GUAN P., HARRIS M.,			
RA	HARRIS N.L., HARVEY D., HEIMLAND T.J., HERNANDEZ J.R., HOUCK J.,			
RA	HOSTIN D., HOUSTON K.A., HOWLAND T.J., WEI M.-H., IBEGWAM C.,			
RA	JALALI M., KALUSH F., KARPEN G.H., KE Z., KENNISON J.A., KETCHUM K.A.,			
RA	KIMMEL B.E., KODIRA C.D., KRAFT C., KRAVITZ S., KULP D., LAI X.,			
RA	LASKO P., LEI Y., LEVITSKY A.A., LI J., LI Z., LIANG Y., LIN Z.,			
RA	LIU X., MATTEI B., MCINTOSH T.C., MCLEOD M.P., MCPHERSON D.,			
RA	MERKULOV G., MILSHINA N.V., MOBARRY C., MORRIS J., MOSHREFI A.,			

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao X., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celnik S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dreenek D., Farfan D.,
RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houch J., Hoskins R.A., Hoslin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirkas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.I., Bregdan R., Bernan B., Carlson J.W., Celnik S.E.,
RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak P., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celnik S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003475; AAF47658.2; -
DR HSSP; Q00534; 1B18.
DR FlyBase; Egn0010909; msn.
DR GO; GO:0000185; P:activation of MAPKKK; NAS.
DR GO; GO:0007391; P:dorsal closure; IMP.
DR GO; GO:0007254; P:JNK cascade; NAS.
DR InterPro; IPR001180; Citron.
DR InterPro; IPR000719; prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TyrKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Transferase.
SQ SEQUENCE 1550 AA; 167219 MW; 780793F89CEB4FC CRC64;

Query Match 40.5%; Score 2805.5; DB 5; Length 1550;
Best Local Similarity 42.2%; Pred. No. 4.1e-153;
Matches 668; Conservative 182; Mismatches 377; Indels 355; Gaps 47;
QY 4 PAPARSLDDIDLSALRDPAGIFELVVEVNGTYQVYKGRHVTKGQLAAIKWMDVTEDEE 63
DB 11 PSVNCSLDDIDLTALKDPAGIFELVVEVNGTYQVYKGRHVTKGQLAAIKWMDVTEDEE 70
QY 64 BEIKOEINMLKYSKSHHNRIATYYGAFIKKSPGNDQDLWMBFCGAGSVTDLVKNTKGN 123
DB 71 BEIKLEINVLKYSKSHHNRIATYYGAFIKKSPGNDQDLWMBFCGAGSVTDLVKNTKGQ 130
QY 124 ALKEDCIAYICREILRGLAHKVIHRDIKGNVLLTENAERVKLVDFGYSALQDRTVG 183
DB 131 SLKEEWIAYICREILRGLSVLHNSKVIHRDIKGNVLLTENAERVKLVDFGYSALQDRTIG 190
QY 184 RRNTFIGHTWMAPEVIACDENPDATYDRSDIWSLIGITALEMAEGAPPLCDMPMRALF 243
DB 191 RRNTFIGHTWMAPEVIACDENPDATYDRSDIWSLIGITALEMAEGAPPLCDMPMRALF 250
QY 244 LIPRNPRLKSKKWSKKFIDFIDTCLIKTVLSRPPTTEOLLKPPFIRODPTEROVRIOLK 303
DB 251 LIPRNPRLKSKKWSKKFIDFIDTCLIKTVLSRPPTTEOLLKPPFIRODPTEROVRIOLK 310
QY 304 DHIDRSKRGKEETEYYSGESEEDSHGEGEPSIMNVPGESTLRREFLRLQENK 363
DB 311 DHIDRCKRKQEKEREDYRSGSDNDDBEPLAGEHSSIVQAPGGDTLRRNFQIQE--- 367
QY 364 SNSALQKQOOLQOQQORDPEAHIKULLHORRIEIQKEERRRVEQQRERERKLOE 423
DB 368 -GRLAEEQOQHHLMAQAQAAAAHAAQAQAQOQQOQAAAAHAAQAQO---AQQ 424
QY 424 KEQORRLDMOALRREERERQAEQEYKRLQEE---ORQSERL----- 465
DB 425 AAQQAQOQOQANQKPRP-----SRQVBEPPPPARPPORLIVVDPHPHANRPLP 475
QY 466 -----QRLQOQEHAYLKSLOOQOQOQOQLQ-----KQOQOQLLP----- 498
DB 476 PTPKCGEPAGTPOQQRNSQNNFKPSLPPRRPEPQPTAAGGGQSQQAQPEAPRNNR 535
QY 499 -----GDRKPL-----YHYGRGNP-----ADKPAWAREVERTR 528
DB 536 QSSGLSSGSGSAGGGGSSKPAALPOQSNHLLGQVNPPLDPSDSDESDEPDPNDRAR 595
QY 529 MNKQNSFLAKSKPGSTGPEP-----PIQASPGPPG-PLSOTPP----- 567
DB 596 -----NDGTLLASDP-----PKPLPGVPVEDANTTTPLSHGSGGPPRPLPPTTDDDDQA 647
QY 568 -----MORVPEQO-----EGPHKS 581
DB 648 GDRTLIMKRLKLEQINRLQKSASTSQANVTPSRRGDSNLLRDWDFRFPKPNANGPRGS 707
QY 582 LQDQPTNLAAPFASH-----DPDPAIPAPT-----ATPSARGAVIRQ- 619
DB 708 GRGSPTTTASLSRSSQLSTLKVDTKLQASVAEAITRPVPRGYQPLKAEPSSASIAKEQ 767
QY 620 -----NSDPTSEGGPSNPMPAVRPDNEAPPKVP-----ORT 652
DB 768 GSASGSGSGSAGSGSGSGSTNSPAHKRQDSDS--RLPTNFERGFPRENSDFFPLAKRY 825
QY 653 SSTATALNTSGA-GGSRPAQAVRARPRNSAWQ---IY-----LQRRAEGRTPKPPGPPAQ 704
DB 826 SAVFSGASASGSGTAGSPSAQALQ---RSSAVYQRNSIYSSSSKSKENDVPAGP-AA 880
QY 705 PPGPPNASSNPDLRRSDPCHWERSDVLPAHGHLPQAGSLERNVGVASSKLDSSPVLSPG 764
DB 881 AAGTATASAKPG-----PAAATTATK-GAAPT----- 908
QY 765 NKAKPDDHRSRPGRPADFVLLKERT-----LDEAPRPKKAMDYSSSSEVESSEDEE 819
DB 909 SKSLANFHLRPR-----EKTESVIVLQNAVPAQRILQLOQQOQQOQQOQQO 960
QY 820 GEGGFAEGSRDTPGGRSDG---DTDSVST-----MVVHD-VVEITGTQPP----- 860


```
QY 819 EGGGPAEGSRDTPGCRSDGDTSDSVTMVVDHVEBITQTQPPYGGTVMVQRTPEERNL 878
Db 635 EG-----NEPLMFKEI 645
QY 879 LHADNGYTNLPDVVQPSHPSTENSKGQSPKSGDGYQSRGLVKAPEKGSF----- 931
Db 646 NASSSRG--ALPDLFPKSPQLRRQINDQTRQMSDDRADEQPNFGFQNSDSRSIQHSFSNR 703
QY 932 ---TWFDVLGIYQPGSGSDSIPTALVGBEGT-----RLDLOLYDVRKGSVVNVNPT 980
Db 704 DREKSFVG---YFGGAG-----AGGTVNRPGRPDQINQVQ-----VNVTN 743
QY 981 N--TRAHSTPIRKYKRFNFSEILCAALGWVNLVGTENGMLLDRSGQKVYGLIGRR 1038
Db 744 SNGTPAENDAPRIKYKRFNFSEILCAALGWVNLVGTENGMLLDRSGQKVYGLIGRR 803
QY 1039 RFQQMDVLEGLNLLTISGRNKLRYVLSWLNKILHND-----PEVEKKQGTWTVGDM 1093
Db 804 RPDQMTVLEGGNILLATISGRKRIRVYVLSWLNKILRTEGAGSANTTEKRGWVNVGDL 863
QY 1094 ECGHYRVVYKRIKFLVIALKSSVEVVAWAPKPYHKEMAFKSFADLPHRLVLDLTVEE 1153
Db 864 QGAIHFIVRYKRIKFLVVGLESEIYAWAPKPYHKEMAFKSFGLSHVPLVLDLTVED 923
QY 1154 GQRLKVIYGSAGFAHVDVDSGNSYDIYIPVHIQSIQITPHAIIFLENTDGMELLCYDE 1213
Db 924 NARLKVLYSGTGFAHDLDAVVDIYTPAQSGQTTTTPHCIVVLPNSNGMQLLCYDNE 983
QY 1214 GYVNTYGRILKDVVLQWGMETPSVAYICSNQIMWGKAKAIBRSVETGHLDPVPMHKA 1273
Db 984 GYVNTYGRMTNVVLQWGMETPSVAYISTGQIMWGKAKAIBRSVDTGHLDPVPMHKA 1043
QY 1274 QRLKFLCERNKDVFPASVRSGSSGVYFMTLNRNCIMNW 1312
Db 1044 QKLKFLCERNKDVFPASSAGGSGCQIYFMTLNLKPLGTNW 1082
RESULT 5
Q9XYC3
ID Q9XYC3 PRELIMINARY; PRT; 1087 AA.
AC Q9XYC3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Serine/threonine protein kinase MIG-15.
GN ZC504.4 OR MIG-15.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Zhu X., Acharya P., Hedgecock E.;
RT "MIG-15, a NIK ortholog of the STR20 family of serine/threonine
RT protein kinases, is involved in cell migration and signal transduction
RT in C. elegans.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kershaw J.K.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF087131; AAD14593.1; -.
DR EMBL; Z50029; CAB63416.1; -.
DR WormPep; ZC504.4a; CE25672.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein.; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001180; Citron.
```

```
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kinase.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00069; kinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1087 AA; 121543 MW; 49FC2C39893F95CD CRC64;

Query Match 38.9%; Score 2696.5; DB 5; Length 1087;
Best Local Similarity 43.7%; Pred. No. Se-1147;
Matches 596; Conservative 157; Mismatches 268; Indels 343; Gaps 34;

QY 10 LDDIDLALRDPAGIFELVEVNGTYGQVYKGRHVKTGQLAAIKVMVDVTEDEEEIKQE 69
Db 6 LDEIDLNSLRDPAGIFELIEVVGNGTYGQVYKGRHVKTGQLAAIKVMVDVTEDEEEIKQE 65
QY 70 INMLKKYSHRNINATYYGATIKKSP--PGNDDQLWLVMEFCGAGSVTDLVKNKTGNALKE 127
Db 66 INMLKKYSHRNINATYYGATIKKSP--PGNDDQLWLVMEFCGAGSVTDLVKNKTGNALKE 125
QY 128 DCIAYICREILRGLAHVHAHVKVIRHDIKQGNVLLTENAELVDFGVSAQLDRTVGRNT 187
Db 126 EWIAIYICREILRGLYHLHQSQVIRHDIKQGNVLLTENAELVDFGVSAQLDRTVGRNT 185
QY 188 FIGPYWMAPEVIACDENPATYDYSIDWSLGTATIAEAGAPLCDMHPMALFLIPR 247
Db 186 FIGPYWMAPEVIACDENPATYDYSIDWSLGTATIAEAGAPLCDMHPMALFLIPR 245
QY 248 NPPRLK-SKWSKKFIDFTCLIKTYLSPPTQQLKPPFIRDOPTQVTRVQLKQHI 306
Db 246 NPPRLKRNKWKTKKFTFETVLVDYHORPYTGALLRHPFIKEQHEQTHIRSIKEHI 305
QY 307 DRSRKKEKEETEYEGSGEEDDSHGEGEPSSIMN-----VPGESTLRREFLRQ 359
Db 306 DNRN--RVKDDADYEYSGSDEDEPSPNNGPSNGIRDDSESSMIAMDNTLRKGFQKQ 363
QY 360 QENKSNSEALKQ--QQLQQQQQRRDPEAHIKHLHQRRRIEEOKE-----ERRRVE 409
Db 364 ESSRGFAEPGAQQLRLPQPAPAP-----FOYQSRVYVPRRESSEVKLRVSSRGAA 417
QY 410 EQQREREQRKLQEKQORLED-----MQALRREEREERQAEQEYKKEQLEEQRS 462
Db 418 DGRPHSPASRPVSVHHQSRPQSHHAPHLADLANYEKRRRSERE----- 463
QY 463 ERLQRLQOQBHAHYLKSLLQQQQQQQLQKQQQLLPQGRKPLVHYGRMNPADKPAWARE 522
Db 464 ERRERERQAHAM----- 476
QY 523 VEERTRNVKQNSPLAKSPGSGTGPBPPIPOASPPGPGPLSQTPPMQRP-----VEQEG 577
Db 477 -----PIARVSASVPAP--QQRKMSSEPLLIITHVKPED- 507
QY 578 PHKSLQDPTNLAAFPASHDPDPAIPATPATPSARGAVIRQNSDPTSEGGPSPNPPAW 637
Db 508 -----LDVLASELSKMGHH-----NGRSRE-----SMSPP-- 534
QY 638 VRPDNEAPPKVPQRTSTATANTSGAGSRPAQAVRARPRNSAWQIY--LQRRABRG 695
Db 535 -----PPAPPREASISITDIDVG-----ELDNGADAEDDLKIMNGE-GT 578
QY 696 KPDPGPPAQPFPNPNASN--PDLRRSDPGWERSDVLDPASHGLHPAGSLERNRVGASS 753
Db 579 LRGENKPL-PPTPTDGTENTLVSDVRNGNG-----NSGH-----GAYKGKKI---- 619
QY 754 KLDSSPVLSPGNKAKPDHRSRPGRPADFVLLKERTLDEAPRPKAMDYSSSESEEVSS 813
Db 620 -----PEIRFGIISLDD----- 634
```



```
KW Hypothetical protein.
SQ SEQUENCE 292 AA; 33574 MW; 768168363DDACC2C CRC64;

Query Match 22.4%; Score 1551; DB 4; Length 292;
Best Local Similarity 100.0%; Pred. No. 1e-81;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1021 MLLDRSGGQKVGGLIGRRRFQMDVLEGLNLLITISGKRNKLRYVYLSWLRNKLILHNDPE 1080
DB 1 MLLDRSGGQKVGGLIGRRRFQMDVLEGLNLLITISGKRNKLRYVYLSWLRNKLILHNDPE 60

QY 1081 VEKQGWTTVGMDEGCHYRVVYKRIKFLVIALKSSVEVYAWAPKPYHKFMAFKSPADL 1140
DB 61 VEKQGWTTVGMDEGCHYRVVYKRIKFLVIALKSSVEVYAWAPKPYHKFMAFKSPADL 120

QY 1141 PHRPILLVLTVEEGQRLKVIYSSAGFAHVDVDSGNSYDIYIPVHIQSOITPHAIIFLPN 1200
DB 121 PHRPILLVLTVEEGQRLKVIYSSAGFAHVDVDSGNSYDIYIPVHIQSOITPHAIIFLPN 180

QY 1201 TDGEMILCYDEGVYNTYGRILIKDVVLQWGMPTSVAYICSNQIMGWGEKAIEIRSV 1260
DB 181 TDGEMILCYDEGVYNTYGRILIKDVVLQWGMPTSVAYICSNQIMGWGEKAIEIRSV 240

QY 1261 TGHLDGVFMHKAQRLKFLCERNDKVFASVRSVSGSSQVYFMTLNRNCIMNW 1312
DB 241 TGHLDGVFMHKAQRLKFLCERNDKVFASVRSVSGSSQVYFMTLNRNCIMNW 292

RESULT 13
QYUAN7 PRELIMINARY; PRT; 792 AA.
AC QYUAN7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Alternatively spliced serine/threonine protein kinase MIG-15
DE (Fragment).
GN MIG-15.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
OC Rhabditidae; Pelodetinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N2;
RA Zhu X., Acharya P., Hedgecock E.;
RT "MIG-15, a NIK ortholog of the STE20 family of serine/threonine
RT protein kinases, is involved in cell migration and signal transduction
RT in C. elegans.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF087132; AAD14594.1; -.
DR PIR; T43630; T43630.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. . .; IEA.
DR InterPro; IPR001180; Citron.
DR Pfam; PF00780; CNH; 1.
DR SMART; SM00036; CNH; 1.
KW Kinase; Serine/threonine-protein kinase.
FT NON TER 1
SQ SEQUENCE 792 AA; 87981 MW; EEE89627662742BE CRC64;

Query Match 21.0%; Score 1458; DB 5; Length 792;
Best Local Similarity 33.9%; Pred. No. 8.9e-76;
Matches 364; Conservative 127; Mismatches 242; Indels 340; Gaps 32;

QY 298 VRIQLKHIDRSRKRGKEETEYYSGSEDDSHGEGEPSSIMN-----VPGEST 350
DB 2 IRRSHKEIDRNR--RVKKDDADYISGSEDDPSFNNRGPFGSMGRDSDSSSSIMPMONT 59

QY 351 LRREFLRLQENKNSKALKKQ--CQLQQQQRRDPEAHIKHLLHQRRIEBOKE----- 403
DB 60 LRKGFQKLQESSRGFAEPGAQQLRLPQPPAP-----FOYQQSRYVEPRRESSEVKL 113
```


Db 241 WSVGITAEMAGAPPLCKLQPLEALCVILREAPKVKSSGWSRKQNFQNMENCMIKNPLF 300

Qy 277 RPPTQLLKFPFIRDPQTERQVRIQLKDH 306

Db 301 RPTSGNMLLHPFVVDIKNERRVVESLTKHL 330

Search completed: August 28, 2004, 01:00:27
Job time : 174 secs

THIS PAGE BLANK (USPTO)


```
Qy 1274 QRLKELCERNKRVFFASVRSGGSSQVYFWTLNRNCIMNW 1312
Db 1037 QRLKELCERNKRVFFSSAKGGSCQIYFWTLNKPGLTNW 1075

RESULT 4
hypothetical protein ZC504.4a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T27622
R:Kershaw, J.
submitted to the EMBL Data Library, July 1995
A:Reference number: Z20394
A:Accession: T27622
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1080 <WIL>
A:Cross-references: EMBL:Z50029; PIDN:CAA90344.1; GSPDB:GN00028; CESP:ZC504.4a
A:Experimental source: clone ZC504
C:Genetics:
A:Gene: CESP:ZC504.4a
A:Map position: X
A:introns: 15/3; 71/3; 137/3; 211/1; 259/1; 388/2; 428/3; 614/3; 643/3; 680/3; 841/3; 95

Query Match 38.2%; Score 2646; DB 2; Length 1080;
Best Local Similarity 43.3%; Pred. No. 1.4e-68;
Matches 590; Conservative 156; Mismatches 268; Indels 350; Gaps 35;

Qy 10 LDDIDLSALRDPAGIFELVEVVGNGTYGQVYGRHVKTGQLAAIKVMQVTEDEEEIKOE 69
Db 6 LDEIDLSLRDPAGIFELIEVVGNGTYGQVYGRHVKTGQLAAIKIMINDEDEDEIKLE 65

Qy 70 INMLKYSHRNIATYGAFFIKSP--PGNDOLWLVMEFCGAGSVTDLVNTKGNALKE 127
Db 66 INMLKXSHRNVAITYGAFFIKLPSSTGKHQDLWLVMEFCGSGSITDLVNTKGGSLKE 125

Qy 128 DCIAYICEILLRGLAHLAHKVIHRDIKGNVLLTENAELVKLVDFGSAQLDRTVGRNT 187
Db 126 EWIAVICREILLRGLYHLQSKVIHRDIKGNVLLTSDAEVLKLVDFGSAQLDRTVGRNT 185

Qy 188 FGTPTWMAPEVIACDENPDATYVRSIDWISLGITAEMAGAPLCPDMHPMALFLIPR 247
Db 186 FGTPTWMAPEVIACDESPATYDSR-----ITALEMAEGHPPLCPDMHPMALFLIPR 238

Qy 248 NPPPLK-SKWSKKFIDFTCLTKLYLSRPPTQLKFPFIRDOPTQVRIQLKDH 306
Db 239 NPPPLKRNKKWKKFETFIETVLVKDYHQRPYTGALLRHPPFKEQHPHQIRHSIKHI 298

Qy 307 DRSRKRGEKETETEYEGSBEEDDSHGEGEPSSIMN-----VPGESTLRREFLRLQ 359
Db 299 DENR--RVKODADYEYSGSEDDPSPNRGPSCGIRDDSESSMIPMDNIRKGFQKLQ 356

Qy 360 QENKNSALKQ-QOQQOQQOORDPEAHKHLHORRIEQQE-----ERRRVE 409
Db 357 ESSRGAEPGAQQLRLPQAPAP-----FQYQSRVYEPRESSEVKLRVSSRGAA 410

Qy 410 EQORERQRKLOEQEQORLED-----MQALRREERQAEQEYKQKLEEQRS 462
Db 411 DGRHPSASRPRPVSHHORSPOQSHAPAPHLADLANYEKKRRSRE----- 456

Qy 463 ERLQRLQOEHAYLKSLQOQQOQQOQLKQOQQOQLPGDKPLVHYGRGMNADKPAPARE 522
Db 457 ERRERQAHAM----- 469

Qy 523 VEERTRMNKQNSPLAKSKPGSTGEPPIPOASPGPGPLSGTPPMQRP-----VBPQEG 577
Db 470 -----PIARVASVPAP-QQSRKMSPLLIITHVKPBD- 500

Qy 578 PHKSLQDQPTRNAAFPAASHDDPAIPAPTATPSARGAVIQNSDPTSEGPGSPNPAP 637
Db 501 -----LDVLAELSCKGGHH-----NGRSREE-----SMSPP-- 527
```

```
Qy 638 VRPDNEAPPKVPQRTSSATATNTSGAGGSRPAQAVRPRSRNSAWQIY--LORRAERGT 695
Db 528 -----PPAPPPREASISSITDIDVG-----ELDNGADAEDWDDLKIMMNGE-GT 571

Qy 696 KPFGPPAQPQPPGNASSN--PDLRSDPQWERSDSVLDPASHGLPOAGSLERNRVGASS 753
Db 572 LRGNPKPL-PPTPTDGGNTLVSDVRNGNG-----NSGH-----GAYKGGKI----- 612

Qy 754 KLDSSPVLSPGNKAKPDHRSRPRGPFADVLLKERTLDEAPRPKKAWDYSSSEEVSS 813
Db 613 -----PEIRPGIISLDD-----DS 627

Qy 814 EDEEEGEGGPAEGSRDTPGGRSDGTDSDSVTMVVDVEEITGTQPPYGGGTVMVQRTPE 873
Db 628 DSDNEEG-----NEPL 638

Qy 874 EERNLLHADSNYTNLPDVVQPSHSPTENSKGQPPSKDGGSDYQSRGLVKAPKSSP-- 931
Db 639 MFKPINASSRSG--ALPDLFPKSPQLRRQINDQTRQMSDDRADQPNQNGFQNSDRSISQH 696

Qy 932 -----TWPVDLGIYQPGSGSDSIPITALVGEGT-----RLDQLOYDVRKGSVV 975
Db 697 SFSNRDREKSPVG--YFGGGAG-----AGGTVNRPRPDQDINQVQ-----V 736

Qy 976 NVNPTN--TRAHSETPIRKYKGFNFSEILCAALWGVNLLVGTENGLMLLDRSGQGVY 1033
Db 737 NVTNSNGTPAENDAPAIRKYKGFSEILCAALWGVNLLIGTDSGLMLLDRSGQGVY 796

Qy 1034 LIGRRRQOMDLBGLNLLITISOKRNKLRYVYLSWLNKILHND-----PEVRKKQGW 1089
Db 797 LISRRRQOMDLBGLNLLITISOKRNKLRYVYLSWLNKILRTEGAGSANTTEKRWGV 856

Qy 1089 TVGMECGCHRVVVKYRIKFLVIALKSSVEVYAWAPKPYHKWAKSPADLPHRPLAVD 1148
Db 857 NVGLQGAIFKIVRYRIKFLVIVVGLSEIEIYAWAPKPYHKWAKSPADLPHRPLAVD 916

Qy 1149 LTVBEGQLKVIYSSAGFAHVDVDSGNSYDIYIPVHIQSIQITPHAIIFLEPNTDGMML 1208
Db 917 LTVEDNARLVLYSGTGGFHAIDLDSAAVVDIYTPAQSGQTTTTHCIVLVFNNGMQLLL 976

Qy 1209 CYDEGVVNTYRIIKDVVLQWGMETSVAYICSNQIMGWGEKAKAIEIRSIVETGLDGVF 1268
Db 977 CYDNEGVVNTYRIIKDVVLQWGMETSVAYICSNQIMGWGEKAKAIEIRSIVETGLDGVF 1036

Qy 1269 MHKRAQRLKFLCERNKRVFFASVRSGGSSQVYFWTLNRNCIMNW 1312
Db 1037 MHKRAQRLKFLCERNKRVFFSSAKGGSCQIYFWTLNKPGLTNW 1080

RESULT 5
T46481
hypothetical protein DKFZp434A025.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T46481
R:Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: 223035
A:Accession: T46481
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1027 <AAA>
A:Cross-references: EMBL:AL137755
A:Experimental source: adult testis; clone DKFZp434A025
C:Genetics:
A:Note: DKFZp434A025.1

Query Match 34.7%; Score 2403.5; DB 2; Length 1027;
Best Local Similarity 51.1%; Pred. No. 1e-61;
Matches 560; Conservative 108; Mismatches 237; Indels 191; Gaps 34;

Qy 262 FIDPIDCLIKTYLSRPPTEQLKFPFIRDOPTQVRIQLKDHDRSKRKGKEET 321
Db 262 FIDPIDCLIKTYLSRPPTEQLKFPFIRDOPTQVRIQLKDHDRSKRKGKEET 321
```


Db 416 EVI---QEN---RYDKVDVWALGVSAIEMAEGLPPRSSVHPNRVLFMISIEPAPMLEDK 470
Qy 257 KWSKKFIDPDTCLIKTYLSRPTTQLLKFPPIRDOPTQVRIOQLKHIDRSRKRKG-- 314
Db 471 KWSLVFHDFAKCLTKPELRPTAAEMLKHKFVERCKTGASA---MSPKIEKSRQIRATM 527
Qy 315 -----EKETEYYSGEEDDSDHGEGEPESSIMNVPGESTLREFLR----- 357
Db 528 ALQAQSVVAPSLEDTSLGPKSSSELGTVPSKPPONSTEAPLTSLNRQHITGNVTLAG 587
Qy 358 -----LQOENKNSSEALKQOQOQLQOQOQORDP 383
Db 588 EGGDFGTWIVHGEDETEESDRSRLVREKSSSSQFEGVP 627

RESULT 9
T34356
hypothetical protein T19A5.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34356
R:Bradshaw, H.
submitted to the EMBL Data Library, March 1996
A:Description: The sequence of C. elegans cosmid T19A5.
A:Reference number: Z21512
A:Accession: T34356
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-653 <BRA>
A:Crosses-references: EMBL:U53153; PIDN:AAC69038.1; GSPDB:GN00023; CESP:T19A5.2
A:Experimental source: strain Bristol N2; clone T19A5
C:Genetics:
A:Gene: CESP:T19A5.2
A:Map position: 5
A:Introns: 26/3; 101/3; 157/1; 209/3; 272/3; 305/2; 373/3; 635/3

Query Match 10.1%; Score 698; DB 2; Length 653;
Best Local Similarity 43.7%; Pred. No. 1.3e-13;
Matches 150; Conservative 64; Mismatches 97; Indels 32; Gaps 9;

Qy 3 DPAPARLDD-----IDLALRDPAIGFELVEVVGNGTYGVYKGRHVKTGOLAAIKVM 56
Db 7 DELPQADDDSMKMDRIYIKL--DDEIFTKQERIGRGSFGEYVGINDRTRGVVAIKII 65

Qy 57 DV--REDEEETKQBINMLKKTSHRNATYYGAFIKKSPGNGDQLMLVMEFCGAGSVT 114
Db 66 DLEQAEDEIEDIQIIVLSQ--CDSQYVTKYFGSLKGS-----KLWIMEYLGGSAL 118

Qy 115 DLVKNKGNALKEDDCIAYICRILGLAHLHAHKVHRDIKQNVLLTENAEVKLVDFGV 174
Db 119 DL---TKSGKLDSEHIAVILREILKGLYLELHSEKIHRIKAANVLVSEHGDKVADFGV 175

Qy 175 SAQLDRTVGRRTFTGTPTWMAPEVITACDENPDATYDYSRDSIWSIGITAIEMAEAPLIC 234
Db 176 AGOLTETVKRITFTGSPFPWMAPELI-----KQSSYDYKADISLIGITAIELANGEPPHS 230

Qy 235 DNHPMRALFLIPNPPRLKSKWKGFIDFTDCLIKTYLSRPTTQLLKFPPIRDOPT 294
Db 231 DLHPNRVFLIPNPPVVLQSQWQSKPFKEFVEMCLNKDPENRPSASTILLKHQFIK-RAK 289

Qy 295 ERQVRIQLKHIDRSRKRKEETEYYSGESEEDDSDHGE 337
Db 290 KNSILVDLIERAAEYRLRTGVS-----SDSLDEDSGGGG 325

RESULT 10
T14157
serine/threonine protein kinase - mouse
C:Species: Mus musculus (house mouse)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14157
R:Pyotowski, B.; Hicklin, D.J.; Kornhaber, G.; Dellaratta, D.V.; Witte, L.

submitted to the EMBL Data Library, December 1997
A:Reference number: Z17894
A:Accession: T14157
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1233 <PVT>
A:Cross-references: EMBL:AF039574; NID:g2773155; PID:g2773156; PIDN:AAB96682.1

Query Match 9.9%; Score 688.5; DB 2; Length 1233;
Best Local Similarity 26.0%; Pred. No. 4.1e-13;
Matches 236; Conservative 145; Mismatches 333; Indels 193; Gaps 37;

Qy 20 DPAGIFELVEVVGNGTYGVYKGRHVKTGOLAAIKVMDVTEDEE-EETKQBINMLKKTYS 78
Db 29 NPBEFWELIGELGDGAFGKVKYAKQNKETNVLAAKVIDTKSEEEEDTWMVIDILAS-CD 87

Qy 79 HRNIATYYGAFIKKSPGNGDQLMLVMEFCGAGSVTDLVKNKGNALKEDCIAYICRIL 138
Db 88 HPNIVKLLDAFY-----YENNLWILIEFCAGGAV-DAVMLELERPLTESQIUVVCKQTL 140

Qy 139 RGLAHLHAHKVHRDIKQNVLLTENAEVKLVDFGSAQLDRTVGRRTFTGTPTWMAPE 198
Db 141 EALNYLDHNKIHRDLKAGNLTLDGDIKLADFGVSAKNTRTIQRRDSFTGTPTWMAPE 200

Qy 199 VIACDENPDATYDYSRDSIWSIGITAIEMAEAPLICDHPMRALFLIPNPPRL-KSKK 257
Db 201 VMCETSKDRPDYDYKADVWSLIGITLIEMAETEPHHELMNRVLLKIAKSPPTLAQPSK 260

Qy 258 WSKKFIDPDTCLIKTYLSRPTTQLLKFPPIRDOPTQVRIOQLKHIDRSRKR----- 313
Db 261 WSSNFKDPLKCKLEKNVDARWTTSQLQHPVT-----VDSNKPVELIA 305

Qy 314 -GEKEETEYYSGESEEDDSDHGEGEPESSIMNVPGESTLREFLR-LRLOQENKNSSEALKQ 371
Db 306 EAKAEVTEEVBDGKEDEEEEAENALP-----IPANKRASSDLSIASSEEDKLSQACIL 360

Qy 372 QOQLQOQOQORDPKAIKH-LLHOR-----QRRTEEQK-----BERRRVEEQ-----QR 413
Db 361 ESVSERTEQSTSEDKFNKILNEKPTTDGPKEKAVDEHASDVNLETGAEINLQTVGIHENG 420

Qy 414 RERQRLQKE--KEQORRLEDQALRREERERQAEQYKRRKQLEEQSERLQRIQOQ 471
Db 421 REKKRPLENLPTDQDQTVDSVSEENNRVLTNTDCLKPEEDRNKE-----NQ 474

Qy 472 EHAYLKSLLQOQO-----QOQOQOQOQOQLLPGDRKPLHYHGRGMNADPAPNAREVEE 525
Db 475 ETLESKLIQSEIINDTHIQTMDLVSQE-----TGEKEADPQAVDN 514

Qy 526 RTRMNKQNSPLAKSKPGSTGPEPPIQASPGPPGLSQTPPMQRPVPEQSGPHKSLQD- 584
Db 515 EVGLTKSET-----QEKLGKDGTAQKVITSDRSSE--VGTDEALDDT 554

Qy 585 QPTRNLAAFPASHDPDPAIPAPTATPSARGAVIRQNSDPTSEGP---GPSNPPPAWVRPD 641
Db 555 QKAELSKAQS GSGDEAL-VPTQT LAEK-----PT-EGPEAGAEPEEPGGERVE 603

Qy 642 NEAP---PKVQRTSSITATALTAGSGSRP-AQAVRAPRPSNSAMQIYLQRRARERTPK 697
Db 604 DKQPEQPAVCEAEAGQLTSTSETTRATLEQDETDEVEQVSESNSIEE--LERLVVTGA-- 659

Qy 698 PPGPPAPPGPPN--ASSNPDLRRSDPCWE---RSDSVLPASHGHLPOAGSLERRNRVGN 752
Db 660 ----BARALGSEGAATAEVDLERKENAQKVFVKAESQAPA-----AS 698

Qy 753 SKLDSSPVLSPGNKAKPDHRSRGRPADFVLLKERTLIDEAPRPKKKAMDYSSSSEEVES 812
Db 699 QPSFPHVPLIPNINSETTENKEMGA--LPKETI-----LPP-----EPEHE 741

Qy 813 SEDDEELEGGAEGSRDTPGGRSDGTD--SVSTMVHVDVEEITGTQPPYGGGTWVORT 871
Db 742 KGNDTDSGTGTVSE-----NSSGDLNLSISSFL-----SKAKDSGVSLOQET 783

Qy 872 PEERNL 878

Db	784	RRQKYL	790	...	
RESULT 11					
S71886					
Ste20-like protein kinase - human					
C:Species: Homo sapiens (man)					
C>Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 21-Jul-2000					
C:Accession: S71886; G01022					
R:Pombo, C.M.; Bonventre, J.V.; Molnar, A.; Kyriakis, J.; Force, T.					
EMBO J. 15, 4537-4546, 1996					
A>Title: Activation of a human Ste20-like kinase by oxidant stress defines a novel stress					
A:Reference number: S71886; MUID:97042345; PMID:8887545					
A:Accession: S71886					
A:Molecule type: mRNA					
A:Residues: 1-426 <POM>					
A:Cross-references: EMBL:X99325; NID:g1430821; PIDN:CAA67700.1; PID:g1430822					
A:Experimental source: cell type B cell					
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog					
C:Keywords: ATP; autophosphorylation; phosphotransferase; protein kinase; stress-induced					
F:18-270/Domain: protein kinase homology <KIN>					
Query Match					
Best Local Similarity 9.8%; Score 678.5; DB 2; Length 426;					
Matches 168; Conservative 68; Mismatches 124; Indels 81; Gaps 15;					
Qy	20	DPAGIFELVEVNGTYGVYKGRHVKTGQLAAIKVMDV--TEDEEIKQEIINMLKYS	77		
Db	15	DPELFTKLDRIKSGFGEVYKIDNHTKEVVAIKIIDLEEAEDSIDIQEITVLSQ-C	73		
Qy	78	HRNRTATYGAFFIKSPGNDQMLVMEFCGAGSVTDLVKNTKGNALKEDCIAVICREI	137		
Db	74	DSPIVTRFGSVLKST-----KLWIMEYLGSSALLDL---KPGPLEETVIATILREI	124		
Qy	138	LRGLAHLAHAKVVIHRDIKQNVLLPENAEVKLVDFGSAQLDRVTGRRNTFTGTPYMAP	197		
Db	125	LKGLDYLHSEKIHRIKAAVLLSEQGVKLDADFGVAGQLTDTQIKRNTFVGTFFWNP	184		
Qy	198	EVIACDENPDATYDYSRWISLIGITAIEMAGAPPLCDHMPRALFLIPRNPPLSKKK	257		
Db	185	EVI-----KQSAIDPKADISLIGITAIEMAGAPPLCDHMPRALFLIPRNPPLSKKK	239		
Qy	258	WSKKFIDITCLIKTYLSRPTEQLKFPPIRQDPTERQVRIQLKHIDRSRKRGEKE	317		
Db	240	SKPKFVEACLNKDFRFTAKELLKHFTIRYTKTSFLTEL---IDRYKR---WKS	292		
Qy	318	ETEYYSGEED-DSHGESEGPSSIMNVGSETLRREFLRLOQENKNSSEALKQQOQLQ	376		
Db	293	EGHGSESSDSIDGEAEDGEQGIWTFP--PTIR-----	326		
Qy	377	QQOQRDPEAHIKHLHQRRRIEEOKEERRRVEEOQRERQ-----RKLOEKEQ	426		
Db	327	-----PSPHSK--LHKGTALHSQKP----ADAVKQPRSQCLSTLVRPVFGELKXKH	374		
Qy	427	Q-----RRLEDQALRRR	439		
Db	375	QSGGSVGALEELNAPSLAE	395		
RESULT 12					
T18532					
serine/threonine protein kinase - guinea pig					
N:Alternate names: STE20-like kinase					
C:Species: Cavia porcellus (guinea pig)					
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000					
C:Accession: T18532					
R:Itch, S.; Kameda, Y.; Yamada, E.; Tsujikawa, K.; Mimura, T.; Kohama, Y.					
Arch. Biochem. Biophys. 340, 201-207, 1997					
A>Title: Molecular cloning and characterization of a novel putative STE20-like kinase in					
A:Reference number: T18952; MUID:97288344; PMID:9143322					
A:Accession: T18532					
A>Status: preliminary; translated from GB/EMBL/DDBJ					
A:Molecule type: mRNA					
A:Residues: 1-1231 <ITO>					
A:Cross-references: EMBL:D88425; NID:g2911113; PIDN:BAA24930.1; PID:g2911114					
A:Experimental source: strain Hartley; liver					
Query Match					
Best Local Similarity 25.1%; Pred. No. 1.4e-12;					
Matches 240; Conservative 166; Mismatches 374; Indels 176; Gaps 36;					
Qy	20	DPAGIFELVEVNGTYGVYKGRHVKTGQLAAIKVMDVTEDEE-BEIKQEIINMLKYS	78		
Db	29	NPEEFWEITIGLGDAGFKVYKQKNTVLAARAVDTKSEBELEDYVMEIDILAS-CD	87		
Qy	79	HRNRTATYGAFFIKSPGNDQMLVMEFCGAGSVTDLVKNTKGNALKEDCIAVICREIL	138		
Db	88	HPNIVKLLDAFY-----YNNMLILIEFCAGGAV-DAVMLELERPLTESQIQVCKQTL	140		
Qy	139	RGLAHLAHAKVVIHRDIKQNVLLPENAEVKLVDFGSAQLDRVTGRRNTFTGTPYMAP	198		
Db	141	EALNYLHDNKKIHRDLKAGNIFLTLDGDIKLDLFGVSAKNTRTTIQRDSFGTTPYMAP	200		
Qy	199	VIACDENPDATYDYSRWISLIGITAIEMAGAPPLCDHMPRALFLIPRNPPLR-KSKK	257		
Db	201	VVMCETSKDRPYDKADVMSLIGITAIEMAEIPEPHHLPNMRVLLKIAKSEPTLAQPSR	260		
Qy	258	WSKKFIDITCLIKTYLSRPTEQLKFPPIRQDPTERQVRIQLKHIDRSRKR---	313		
Db	261	WSSNFKDPLKCKENVDARTTSQLLOHPVT-----IDSNKPIRELIJA	305		
Qy	314	-GEKETETEYVSGSEEDDSHGESEP-----SSIMNVG--ESTLRR-----EFLRL	358		
Db	306	EAKAEVTEVEDGKEEDDEEISLPIPTNKRASSDLSSASSEDKLSONACILESVSE	365		
Qy	359	QOENKNSSEALKQQOQLQOQQOORDEPAHIK-----HLLHQRRRIEEOKEERRRVEEQ	412		
Db	366	KTEHNASGDKFSTKVLNKEPCGEPENAVELVGAVAVLPDRATELPESGREERKPLDR	425		
Qy	413	RREROQKLOEKOORLEDMQALRR--EERQAEEOEYKQKLEOR--QSERLQR	467		
Db	426	LPOTDEDEMADINSVSEGEDHATVSTNIEHNLKPKERDOEKQPVLENKLVKSEDTTI	485		
Qy	468	Q-----LOQE-----HAYLSLQOQQOQQOQLQOQQOQLLPGRKPLVHY	507		
Db	486	QTVDLVSOETGEKVDTHILDSEVVHAV-----EDTHEKLKDDTTQ-----KDLV--	531		
Qy	508	GRGNVPADKAWAREVEERTMKNQNSPLAKSPGSTGPEPP-----IPQASPGPGP	561		
Db	532	-----SDTSVGRDEIEIGAVPKTAESS-AEGAQGGGKETDEGAQILISKATEGPKAS	584		
Qy	562	-LSOTPPMQRPVPEQEPHKSLODPTNLAAFPASHDDPDAIPAPTATPSARGAVIRON	620		
Db	585	GTEEAPPVTEITENDTDQLVEN-----THEKQLPISSETLDTSEGLHASEG	633		
Qy	621	SDPTSEGFSPNPPAWVRPDNEAPPKVPQRTSSIA-TALNTSGAGGSRPAQAVRPR	679		
Db	634	REVTESGSTEVEVEGAVSETDEEDVQSETRGAPMAVTQMDTEKNETPHEAPA-----	686		
Qy	680	NSAQIYLQRAERGTKPPGPAQPPGPP-PNASSNPDL--RRSDPGWE-RSDSVLPASH	735		
Db	687	----QVEVQ-----VPVPPQSPGPPAPIPSPININSEAAENKGMGASLNTETIL----	732		
Qy	736	GHLPQAGSLERNV--GASSKLDSSPV---LSPGN-KAKPDDHRS---RGRPADFVLK	786		
Db	733	--LPSESSQKENDTDSGTSTADNSSLIDNLSSFLSKTKYDNGSISLQETRRQKTKLK	790		
Qy	787	ERT-----LDEAPRPPKAMDYSSSSEE-----VESSEDDDEEGGPAEG	827		
Db	791	TRKPIVDGVSVTTSKIIVTDSKTEELRFLRRQELRELFLQKEQRAQQQLNGKLQQ	850		
Qy	828	SRDTPGSRGSDTDSVSTVMVHVEITGTGTPPYGGTGMVQVTPPEERNLLHADS	883		
Db	851	QREQIFRRFEQEMMSKKRQYDOEINLEKQO-----KQTIERLEQSHTRNLRDEA	900		

QY 298 VRIQLKDHIDRSKRKRGEETEYE-----YSG-----SEEDDSHG----- 334
 Db 287 -ETVLIDLITQKDAVDRLNLDQYRKMKLLFOEAHNGPVAEQAEEHEEDHGGGRTGTV 345
 QY 335 -----EPEGSSIMNVPGESTLRREFLRLOQEN--KSNSEA--LKQOQOOL 375
 Db 346 NSVGNQSPISMSISASSQSSVNSLPDASDDKSELDMWEGDHTVMSNSSVTHLKPEEE- 404
 QY 376 QOQOQORDPEAH-----KULLHOR-----RIEQKEERRRVEEQ- 411
 Db 405 NYQEEGDPRTRASAPQSPQVSRHKSHYRNREHFATIRTSALVTROMQEHQDSSELREQM 464
 QY 412 ----QRREREQKLOEQEORLE--DMOALRREE-----RROAEREQ 449
 Db 465 SOYKEMRQHOKQMTLENKLAENDEHRLDKDLETQRNFAEMEKLIKQHASMEK 524
 QY 450 EYRKOLBEQROSERLQROLOE--HAYLKS-----LOOQOQOQOOLQ-----KQOQOOL 496
 Db 525 EAKVMAEBEKKPQHQIQAQKKELNSFLSQKREYKLAKEQLKEELNENQSTPKKEKQEW 584
 QY 497 LPQDRKPLHYGRGNPNADKPAAWREVEERTMKNQNSPL 537
 Db 585 LSKQKENTQHF-----QAEENALLRRORQYL 611

 RESULT 15
 AS3714
 protein kinase (EC 2.7.1.37) BL44 - human
 N:Alternate names: GC kinase
 C:Species: Homo sapiens (man)
 C:Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 24-Sep-1999
 R:Katz, P.; Whalen, G.; Kehrl, J.H.
 J. Biol. Chem. 269, 16802-16809, 1994
 A:Title: Differential expression of a novel protein kinase in human B lymphocytes. Prefe
 A:Reference number: AS3714; MUID:94266900; PMID:7515885
 A:Accession: AS3714
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-819 <KAT>
 A:Cross-references: GB:U07349; NID:g531819; PIDN:AAA20968.1; PID:g531820
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
 C:Keywords: ATP; phosphotransferase
 F:13-272/Domain; protein kinase homology <KIN>
 F:21-29/Region; protein kinase ATP-binding motif

 Query Match 9.4%; Score 653.5; DB 2; Length 819;
 Best Local Similarity 29.8%; Pred. No. 2.9e-12;
 Matches 186; Conservative 84; Mismatches 181; Indels 173; Gaps 23;

 QY 17 ALRDPAGIFELVEVNGTYGVYKGRHVKTQOLAIAIKVMDVTE--DEEEIKQEIINLKK 75
 Db 7 SLQDPDRFELLQVRGTYGVYKARDVTVTSELAAVKIVKLDPGDDISSIQQEITILRE 66

 QY 76 YSHRNIAITYGAFIKKSPGNDQDLWLMFPCGAGSVTDLVKNTKGNALKEDCIAYICR 135
 Db 67 -CRHFNVAIYGSYL-----NDRLWICMEFCGGSLSQEIYHAT--GPLEERQIAYVCR 117

 QY 136 EILRGLAHLHAHVTHRIDKQNVLLTENAIEVKLVDFGVSQQLDRVTGRRNTFTGTPYWM 195
 Db 118 ERLGLHLHLSQCKTHRIDKGANLLLTQGDVKLADFGVSGELTASVAKRRSFTGTPYWM 177

 QY 196 APEVTACDENPDATYDYSIDTSLGITAIEAEGAPPLCDMHPMALEFLIPN--PPRL 253
 Db 178 APEVAVERK--GGYNEICDVMALGITAIELEGLOPFLHLHPMALMLMSKSSFPKPL 235

 QY 254 KSK-KWSKKFIDFIDTCLIKTYLSRPPTEQLIKFPFIRDQDPTERQVRIQLKDHIDRSRKX 312
 Db 236 RDKTRWTQNFHFLKALTKNPKKPTAEKLLQHPFTTQQ-LPRALLTQLLDKASDPHLG 294

 QY 313 RGEETEYE-YSGSEEDDSHGEGEPSSIMNVPGESTLRREFLRLOQENKSNSEALKQ 371

Db 295 TPSPEDCELETYDMFPDTIHSRGQH-----PAERT----- 325
 QY 372 QOQLQOQOQORDPEAHIKHLLHORRRITEEQKEERRRVEEQRREREQKLOEKQORRELE 431
 Db 326 PSEIQ-----FHQ-----VKFGAPRRKETDPLNEPWE-----E 353
 QY 432 DMQALRREERERRQAEREQEVKQKLEBQROSERLQRLQQHAYLKSLOOQOQOQOQKQ 491
 Db 354 EWTLLGKEE-----LGSLLQSVQEALEERSLTIR 383
 QY 492 QOQOQLLPGDRKPLHYGRGNPNADKPAAWREVEERTMKNQNSPLAKSKFGSGTGPPEPI 551
 Db 384 SASEFQBLD-----SPDD-----TMGTIKRAPFL-----GPLFTDPPA 416
 QY 552 POASGPPGGLSOTPPMQRPVPEQEGPHKSLQDQPTRNLAAFPASHDDPDAIPAPTATPS 611
 Db 417 EEPUSPPGTU-----PP-----PPSGPNS-----PLLTAWAT-- 446
 QY 612 ARGAVIRQNSDP-----TSEGGPQSP 632
 Db 447 -----MKQREDPERSSCHGLPPTP 465

Search completed: August 28, 2004, 01:01:22
 Job time : 65 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 28, 2004, 00:57:10 ; Search time 81 Seconds
(without alignments)
5095.952 Million cell updates/sec

Title: US-10-029-115-2
Perfect score: 6929
Sequence: 1 MGDPAARSLLDIDLSALRD.....SGSSQVYFMTLNRNCIMNW 1312

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1297172 seqs, 314612898 residues

Total number of hits satisfying chosen parameters: 1297172

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6929	100.0	1312	14	US-10-029-115-2
2	6660.5	96.1	1295	10	US-09-789-390-30
3	6660.5	96.1	1295	10	US-09-789-390-32
4	6660.5	96.1	1295	10	US-09-789-390-34
5	6660.5	96.1	1295	10	US-09-789-390-37
6	6660.5	96.1	1295	10	US-09-789-390-39
7	6659	95.1	1326	10	US-09-291-417-15
8	6646.5	95.9	1303	10	US-09-789-390-35
9	6646.5	95.9	1303	10	US-09-789-390-38
10	6628.5	95.7	1303	10	US-09-789-390-9
11	6622	95.6	1332	10	US-09-789-390-7
12	6603.5	95.3	1276	14	US-10-029-115-6
13	6311	91.1	1244	10	US-09-789-390-13
14	6304.5	91.0	1273	10	US-09-789-390-11
15	5422.5	78.3	1303	10	US-09-789-390-65

16	5412	78.1	1332	10	US-09-789-390-63	Sequence 63, Appl
17	5261	75.9	1244	10	US-09-789-390-70	Sequence 70, Appl
18	5261	75.9	1244	10	US-09-789-390-71	Sequence 71, Appl
19	5250.5	75.8	1273	10	US-09-789-390-67	Sequence 67, Appl
20	5250.5	75.8	1273	10	US-09-789-390-68	Sequence 68, Appl
21	4336	62.6	1360	9	US-09-871-916-2	Sequence 2, Appli
22	4336	62.6	1360	14	US-10-355-975-14	Sequence 14, Appl
23	4274	61.7	1268	15	US-10-353-690-122	Sequence 122, App
24	4195.5	60.5	1297	10	US-09-291-417-14	Sequence 14, Appl
25	4021.5	58.0	1239	10	US-09-291-417-13	Sequence 13, Appl
26	3977	57.4	1212	12	US-10-168-582-9	Sequence 9, Appli
27	3977	57.4	1212	14	US-10-247-671-157	Sequence 157, App
28	3951.5	57.0	1233	10	US-09-291-417-89	Sequence 89, Appl
29	3872.5	55.9	1165	12	US-10-211-462-89	Sequence 89, Appl
30	3872.5	55.9	1165	14	US-10-021-660-126	Sequence 126, App
31	3824	55.2	1175	9	US-09-771-161A-224	Sequence 224, App
32	3824	55.2	1175	9	US-09-771-161A-225	Sequence 225, App
33	3824	55.2	1175	9	US-09-771-161A-226	Sequence 226, App
34	3821	55.1	792	14	US-10-029-115-4	Sequence 4, Appli
35	2708.5	39.1	1109	10	US-09-291-417-88	Sequence 88, Appl
36	2403.5	34.7	1027	16	US-10-408-765A-1181	Sequence 1181, Ap
37	2280	32.9	510	9	US-09-925-297-596	Sequence 596, App
38	1862.5	26.9	1227	10	US-09-291-417-105	Sequence 105, App
39	1819.5	26.3	1581	12	US-10-415-011-16	Sequence 16, Appl
40	1797	25.9	515	9	US-09-771-161A-134	Sequence 134, App
41	1686	24.3	425	9	US-09-771-161A-133	Sequence 133, App
42	1676.5	24.2	468	9	US-09-771-161A-135	Sequence 135, App
43	1600.5	23.1	339	16	US-10-664-421-79	Sequence 79, Appl
44	1532	22.1	288	12	US-10-276-1774-1920	Sequence 1920, Ap
45	1457	21.0	275	14	US-10-355-975-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-10-029-115-2
; Sequence 2, Application US/10029115
; Publication No. US20030077597A1
; GENERAL INFORMATION:
; APPLICANT: Luo, Ying
; APPLICANT: Fu, Alan C
; APPLICANT: Shen, Mary
; TITLE OF INVENTION: No. US20030077597A1el Germlinal Center Kinase Cell Cycle Protiens,
; FILE REFERENCE: A-70229/RMS/DHR
; CURRENT APPLICATION NUMBER: US/10/029,115
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 1312
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-115-2

Query Match	100.0%;	Score	6929;	DB	14;	Length	1312;
Best Local Similarity	100.0%;	Pred. No.	4.4e-311;				
Matches	1312;	Conservative	0;	Mismatches	0;	Indels	0;
		Gaps	0;				
Qy	1	MGDPAPARSLLDIDLSALRDPAAGIFELVEVVGNGTGYGVYKGRHVKTGQLAAIKVMDVTE	60				
Db	1	MGDPAPARSLLDIDLSALRDPAAGIFELVEVVGNGTGYGVYKGRHVKTGQLAAIKVMDVTE	60				
Qy	61	DEEEIKQEIINMLKKYSHHRNIATYYGAFIKKSPGNDQDLWLVMFCGAGSVTDLVKNT	120				
Db	61	DEEEIKQEIINMLKKYSHHRNIATYYGAFIKKSPGNDQDLWLVMFCGAGSVTDLVKNT	120				
Qy	121	KGNAKEDCTIAYICREILRGLAHLHAHKVIRHDIKQGNVLLTENAELVKLVDFGVSQAQLDR	180				
Db	121	KGNAKEDCTIAYICREILRGLAHLHAHKVIRHDIKQGNVLLTENAELVKLVDFGVSQAQLDR	180				
Qy	181	TVGRNFTFGTPYMWAEVACDENPDATYDYSGLGITAEAGAPPLCDMHPMR	240				

Db 481 QOQQOQLOKQOQOQLPGDRKPLHYHGRGMNPKAPAWAREVEERTMKNQONSPLAKS 540
Qy 541 KPGSTGPEPPIPOASGPGPGPLSQTTPMQRVPVEQEGPHK----- 580
Db 541 KPGSTGPEPPIPOASGPGPGPLSQTTPMQRVPVEQEGPHKSLVAHRVPLKPYAAPVRSQ 600
Qy 581 SLQDQPTNLAAFPASHDDPAIPAPTATPSARGAVIRONSDDTSEGPGSPNPAAWRP 640
Db 601 SLQDQPTNLAAFPASHDDPAIPAPTATPSARGAVIRONSDDTSEGPGSPNPAAWRP 660
Qy 641 DNEAPPKVPQRTSSIALNTALNTSGAGSRPAQAVRPRSNASAWQIYLORRAERTGPKPPG 700
Db 661 DNEAPPKVPQRTSSIALNTALNTSGAGSRPAQAVRA----- 695
Qy 701 PPAQPPGPNASNDLRRSDPGWERSDVLPAASHGLPQAGSLERNRVGASSKLDSPV 760
Db 696 -----SNPDLRSDPGWERSDVLPAASHGLPQAGSLERNRVGSSKPDSPV 743
Qy 761 LSPGNKAPDDHRSRPGPADPVLLKERTLDEAPPPKAMDYSSSEVESSEDDDEEG 820
Db 744 LSPGNKAPDDHRSRPGPADPVLLKERTLDEAPPPKAMDYSSSEVESSEDDDEEG 803
Qy 821 EGGPAEGSRDTPGGRSDGDTDSVSTWVHDVBEITGTQPPYGGGTWVQRTPEERNLLH 880
Db 804 EGGPAEGSRDTPGGRSDGDTDSVSTWVHDVBEITGTQPPYGGGTWVQRTPEERNLLH 863
Qy 881 ADSNGYTNLPDVVQPSHSPTENSKQSPSKDGSQYSGRLVAPKXSSFTMFVDLGIY 940
Db 864 ADSNGYTNLPDVVQPSHSPTENSKQSPSKDGSQYSGRLVAPKXSSFTMFVDLGIY 923
Qy 941 QPGSGSDSIPITALVGGEGTRLDQLOYDVYKGSVNVNPTNTRAHSETPEIRKYKRFNS 1000
Db 924 QPGSGSDSIPITALVGGEGTRLDQLOYDVYKGSVNVNPTNTRAHSETPEIRKYKRFNS 983
Qy 1001 ETLCAALMGVNLVGTENGLMLDRSGGQKVGILGRRRFOQMDVLEGLNLLITISGRN 1060
Db 984 ETLCAALMGVNLVGTENGLMLDRSGGQKVGILGRRRFOQMDVLEGLNLLITISGRN 1043
Qy 1061 KLRVYLSWLRNKLHNDPEVEKKQGTWTVGDMGCGHYRVVYKRIKFLVIALKSSVEV 1120
Db 1044 KLRVYLSWLRNKLHNDPEVEKKQGTWTVGDMGCGHYRVVYKRIKFLVIALKSSVEV 1103
Qy 1121 YAWAPKPHKFAFKSPADLPHRPLLDLTVEEGORLKVYIGSSAGFAHVDVDSGNSYDI 1180
Db 1104 YAWAPKPHKFAFKSPADLPHRPLLDLTVEEGORLKVYIGSSAGFAHVDVDSGNSYDI 1163
Qy 1181 YIPVHIQSOITPHALIFLPTDGMELLCYEDEGVVYNTYGRILKDVVLQWCEMPTSVAY 1240
Db 1164 YIPVHIQSOITPHALIFLPTDGMELLCYEDEGVVYNTYGRILKDVVLQWCEMPTSVAY 1223
Qy 1241 ICNQIMGWGEKAIBIRSVETGHLDPGMHKAQRLKFLCERNDKVFFASVRSQSSQVY 1300
Db 1224 ICNQIMGWGEKAIBIRSVETGHLDPGMHKAQRLKFLCERNDKVFFASVRSQSSQVY 1283
Qy 1301 FMTLNRNCIMNW 1312
Db 1284 FMTLNRNCIMNW 1295

RESULT 3

US-09-789-390-32
; Sequence 32, Application US/09789390
; Publication No. US20030059768A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: MacDougall, John
; APPLICANT: Shinkets, Richard A
; APPLICANT: Spaderna, Steven K
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-692
; CURRENT APPLICATION NUMBER: US/09/789,390
; CURRENT FILING DATE: 2001-02-23

; PRIOR APPLICATION NUMBER: 60/185,548
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/199,957
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/184,951
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 60/185,967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/197,723
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 1295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-789-390-32

Query Match 96.1%; Score 6660.5; DB 10; Length 1295;
Best Local Similarity 95.6%; Pred. No. 1e-298;
Matches 1273; Conservative 0; Mismatches 2; Indels 57; Gaps 2;

Qy 1 MGDPAARSDDIDLSALRDPAGIFELVEVVGNGTYGOVYKGRHVKTGQLAAIKVMDVTE 60
Db 1 MGDPAARSDDIDLSALRDPAGIFELVEVVGNGTYGOVYKGRHVKTGQLAAIKVMDVTE 60
Qy 61 DEEBEIKQEIINMLKYSHRRNIATYYGAFIKKSPGNDQDLWLVMEFCGAGSVTDLVKNT 120
Db 61 DEEBEIKQEIINMLKYSHRRNIATYYGAFIKKSPGNDQDLWLVMEFCGAGSVTDLVKNT 120
Qy 121 KGNALKECDIAYICREILRGLAHHAHVIRHDIKQNVLLTENAENVKLVDFGVSQAQLDR 180
Db 121 KGNALKECDIAYICREILRGLAHHAHVIRHDIKQNVLLTENAENVKLVDFGVSQAQLDR 180
Qy 181 TVGRNRTFIGTPYMAVEVIACDENPDATDYRSDIWSLGITALEMAGAPPLCDMHPMR 240
Db 181 TVGRNRTFIGTPYMAVEVIACDENPDATDYRSDIWSLGITALEMAGAPPLCDMHPMR 240
Qy 241 ALFLIPRNPRLSKSKWKKFIDFIDTCLIKTYLSRPPTEQLLKFPPIRQPTERQVRI 300
Db 241 ALFLIPRNPRLSKSKWKKFIDFIDTCLIKTYLSRPPTEQLLKFPPIRQPTERQVRI 300
Qy 301 QLKDHIDRSKRKGEKETEYEGSSEEDDSDHGECEGEPSSIMNVPGESTLRREFLRLQO 360
Db 301 QLKDHIDRSKRKGEKETEYEGSSEEDDSDHGECEGEPSSIMNVPGESTLRREFLRLQO 360
Qy 361 ENKNSBALKQOQLOQOQOQDPEAHIKHLHQRRRIEOKERRRVEEQRREREORQK 420
Db 361 ENKNSBALKQOQLOQOQOQDPEAHIKHLHQRRRIEOKERRRVEEQRREREORQK 420
Qy 421 LQEKQOQRLLEDMAALREERROAEREQYKRLQLEEQRQSERLQRLQOQEHAYLSLQ 480
Db 421 LQEKQOQRLLEDMAALREERROAEREQYKRLQLEEQRQSERLQRLQOQEHAYLSLQ 480
Qy 481 QOQOQOQLOQOQOQOQLPGDRKPLHYHGRGMNPKAPAWAREVEERTMKNQONSPLAKS 540
Db 481 QOQOQOQLOQOQOQOQLPGDRKPLHYHGRGMNPKAPAWAREVEERTMKNQONSPLAKS 540
Qy 541 KPGSTGPEPPIPOASGPGPGPLSQTTPMQRVPVEQEGPHKSLVAHRVPLKPYAAPVRSQ 600
Db 541 KPGSTGPEPPIPOASGPGPGPLSQTTPMQRVPVEQEGPHKSLVAHRVPLKPYAAPVRSQ 600
Qy 581 SLQDQPTNLAAFPASHDDPAIPAPTATPSARGAVIRONSDDTSEGPGSPNPAAWRP 640
Db 601 SLQDQPTNLAAFPASHDDPAIPAPTATPSARGAVIRONSDDTSEGPGSPNPAAWRP 660
Qy 641 DNEAPPKVPQRTSSIALNTALNTSGAGSRPAQAVRPRSNASAWQIYLORRAERTGPKPPG 700
Db 661 DNEAPPKVPQRTSSIALNTALNTSGAGSRPAQAVRA----- 695
Qy 701 PPAQPPGPNASNDLRRSDPGWERSDVLPAASHGLPQAGSLERNRVGASSKLDSPV 760
Db 696 -----SNPDLRSDPGWERSDVLPAASHGLPQAGSLERNRVGSSKPDSPV 743


```
QY 1001 EILCAALMGVNLVGTENGMLLDRSGGQKVYGLIGRRRFOQMDVLEGLNLLITISGRN 1060
DB 984 EILCAALMGVNLVGTENGMLLDRSGGQKVYGLIGRRRFOQMDVLEGLNLLITISGRN 1043
QY 1061 KLRVYLSLNRNKILHNDPEVEKKQGWTTVGMGCGHYRVVVKYERIKFLVIALKSSVEV 1120
DB 1044 KLRVYLSLNRNKILHNDPEVEKKQGWTTVGMGCGHYRVVVKYERIKFLVIALKSSVEV 1103
QY 1121 YAWAPKPYHKFMFAPSADLPHRPLLVDLTVBEGQRLKVIYSSAGFAHVDVDSGNSYDI 1180
DB 1104 YAWAPKPYHKFMFAPSADLPHRPLLVDLTVBEGQRLKVIYSSAGFAHVDVDSGNSYDI 1163
QY 1181 YIPVHIQSQTTHAIIFLPNTDGMELLCYEDEGVVNTYGRILKDVVLQWGMPTSVAY 1240
DB 1164 YIPVHIQSQTTHAIIFLPNTDGMELLCYEDEGVVNTYGRILKDVVLQWGMPTSVAY 1223
QY 1241 ICSNQIMGWGEKAIEIRSVETGHLDGVMFKHAQRLKFLCERNDKVFFASVRSQSSQVY 1300
DB 1224 ICSNQIMGWGEKAIEIRSVETGHLDGVMFKHAQRLKFLCERNDKVFFASVRSQSSQVY 1283
QY 1301 FMTLNRNCIMNW 1312
DB 1284 FMTLNRNCIMNW 1295
```

RESULT 5

```
US-09-789-390-37
; Sequence 37, Application US/09789390
; Publication No. US20030059768A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: MacDougall, John
; APPLICANT: Shimkets, Richard A
; APPLICANT: Spaderna, Steven K
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-692
; CURRENT APPLICATION NUMBER: US/09/789,390
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185,548
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/199,957
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/184,951
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 60/185,967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/197,723
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 1295
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-789-390-37
```

```
Query Match 96.1%; Score 6660.5; DB 10; Length 1295;
Best Local Similarity 95.6%; Pred. No. 1e-298;
Matches 1273; Conservative 0; Mismatches 2; Indels 57; Gaps 2;

QY 1 MGDPAARSLLDDIDLSALRDPAGIFELVEVVGNGTYGVYGRHVKVTGQLAAIKVMDVTE 60
DB 1 MGDPAARSLLDDIDLSALRDPAGIFELVEVVGNGTYGVYGRHVKVTGQLAAIKVMDVTE 60
QY 61 DEEEIKQEIINMLKKYSHRNNTATYGAFFIKSPGNDLQWLVMFCGAGSVTLVKNT 120
DB 61 DEEEIKQEIINMLKKYSHRNNTATYGAFFIKSPGNDLQWLVMFCGAGSVTLVKNT 120
QY 121 KGNALKEDCIAYICREILRGLAHLHAHKVIHRDIKQNVLLTENAELVKLVDFGVAQLDR 180
DB 121 KGNALKEDCIAYICREILRGLAHLHAHKVIHRDIKQNVLLTENAELVKLVDFGVAQLDR 180
```

```
QY 181 TVGRNTEIGTYWMAPEVIAICDENPDATYDYSNDIWSLGITALEMAGAPPLCDHMPWR 240
DB 181 TVGRNTEIGTYWMAPEVIAICDENPDATYDYSNDIWSLGITALEMAGAPPLCDHMPWR 240
QY 241 ALFLIPRNPPLKSKWKKFIDFIDTCLTKTYLSRPPTEQLLKFPPIRDOPTERQVRI 300
DB 241 ALFLIPRNPPLKSKWKKFIDFIDTCLTKTYLSRPPTEQLLKFPPIRDOPTERQVRI 300
QY 301 QLKDHIDRSRKRGKEKETETEYSGSEBEDSHGEGEPSSIMNVPGSESTLRREFLRLOQ 360
DB 301 QLKDHIDRSRKRGKEKETETEYSGSEBEDSHGEGEPSSIMNVPGSESTLRREFLRLOQ 360
QY 361 ENKNSSEALKQOQOQLQOQOQORDPEAHIKHLHQRRIIEQEKERRRRVEEQORRREROK 420
DB 361 ENKNSSEALKQOQOQLQOQOQORDPEAHIKHLHQRRIIEQEKERRRRVEEQORRREROK 420
QY 421 LOEKEQORRLDMQALRREERROAEREQEKYRQLEEQORQSERLQRLQOQEHAYLKSILQ 480
DB 421 LOEKEQORRLDMQALRREERROAEREQEKYRQLEEQORQSERLQRLQOQEHAYLKSILQ 480
QY 481 QOQOQOQLQOQOQOQLLPQDRKPLYHYGRGNWPKADKPAWAREVEERTMKNQONSPLAKS 540
DB 481 QOQOQOQLQOQOQOQLLPQDRKPLYHYGRGNWPKADKPAWAREVEERTMKNQONSPLAKS 540
QY 541 KPGSTGPEPPIPOASPGPGPLSOTPPMORPVEPQEGPHK----- 580
DB 541 KPGSTGPEPPIPOASPGPGPLSOTPPMORPVEPQEGPHKSLVAHRVLPKYAAPVPRSQ 600
QY 581 SLOQOPTNLAAFPASHDPDPAPATPATPSARGAVIRQNSDPTSEGGPGSPNPPAWVRP 640
DB 601 SLOQOPTNLAAFPASHDPDPAPATPATPSARGAVIRQNSDPTSEGGPGSPNPPAWVRP 660
QY 641 DNEAPPKVPQRTSSIALNTSGAGGSRPAQAVRARPRSNSAWQIYLQRRABRGTPKPPG 700
DB 661 DNEAPPKVPQRTSSIALNTSGAGGSRPAQAVRA----- 695
QY 701 PPAQPGPPNASSNPDLRRSDPGHERSDVLPASHGHLPOAGSLERNRVGASSKLDSPV 760
DB 696 -----SNPDLRRSDPGHERSDVLPASHGHLPOAGSLERNRVGASSKLDSPV 743
QY 761 LSPGNKAKPDDHRSRPGRPADFVLLKERTLDEAPRPKAMDYSSSSSEVESSEDDDEEG 820
DB 744 LSPGNKAKPDDHRSRPGRPADFVLLKERTLDEAPRPKAMDYSSSSSEVESSEDDDEEG 803
QY 821 EGGPAEGSRDTPGGRSDGTDSDVSTVMVVDVEEITGTQPPYGGGTVMVQRTPEBERNLLH 880
DB 804 EGGPAEGSRDTPGGRSDGTDSDVSTVMVVDVEEITGTQPPYGGGTVMVQRTPEBERNLLH 863
QY 881 ADSNGYTNLPDVVQPSHSPTEKSGQSPKSGDYGQSRGLVKAPGKSSFTMFVDLGIY 940
DB 864 ADSNGYTNLPDVVQPSHSPTEKSGQSPKSGDYGQSRGLVKAPGKSSFTMFVDLGIY 923
QY 941 QPGSGSDSIPITALVGGEGTRLDQLOQYDVRKGSVVNVNPTNTRAHSETPEIRKYKRFNS 1000
DB 924 QPGSGSDSIPITALVGGEGTRLDQLOQYDVRKGSVVNVNPTNTRAHSETPEIRKYKRFNS 983
QY 1001 EILCAALMGVNLVGTENGMLLDRSGGQKVYGLIGRRRFOQMDVLEGLNLLITISGRN 1060
DB 984 EILCAALMGVNLVGTENGMLLDRSGGQKVYGLIGRRRFOQMDVLEGLNLLITISGRN 1043
QY 1061 KLRVYLSLNRNKILHNDPEVEKKQGWTTVGMGCGHYRVVVKYERIKFLVIALKSSVEV 1120
DB 1044 KLRVYLSLNRNKILHNDPEVEKKQGWTTVGMGCGHYRVVVKYERIKFLVIALKSSVEV 1103
QY 1121 YAWAPKPYHKFMFAPSADLPHRPLLVDLTVBEGQRLKVIYSSAGFAHVDVDSGNSYDI 1180
DB 1104 YAWAPKPYHKFMFAPSADLPHRPLLVDLTVBEGQRLKVIYSSAGFAHVDVDSGNSYDI 1163
QY 1181 YIPVHIQSQTTHAIIFLPNTDGMELLCYEDEGVVNTYGRILKDVVLQWGMPTSVAY 1240
DB 1164 YIPVHIQSQTTHAIIFLPNTDGMELLCYEDEGVVNTYGRILKDVVLQWGMPTSVAY 1223
QY 1241 ICSNQIMGWGEKAIEIRSVETGHLDGVMFKHAQRLKFLCERNDKVFFASVRSQSSQVY 1300
```


[illegible]


```
; FILE REFERENCE: 15966-692
; CURRENT APPLICATION NUMBER: US/09/789,390
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185,548
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/199,957
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/184,951
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 60/185,967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/197,723
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 1273
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-789-390-11

Query Match          91.0%; Score 6304.5; DB 10; Length 1273;
Best Local Similarity 90.5%; Pred. No. 2.5e-282;
Matches 1221; Conservative 2; Mismatches 13; Indels 113; Gaps 5;

Qy 1 MGDPAARSDDIDLSALRDPAGIPELVEVVGNGTYGQYKGRHVKTGQLAAIKVMVTE 60
Db 1 MGDPAARSDDIDLSALRDPAGIPELVEVVGNGTYGQYKGRHVKTGQLAAIKVMVTE 60

Qy 61 DEEEIKOEINMKKYSHNRNIATYGAFFKSPGNDQDLWMEFCGAGSVTDLVKNT 120
Db 61 DEEEIKOEINMKKYSHNRNIATYGAFFKSPGNDQDLWMEFCGAGSVTDLVKNT 120

Qy 121 KGNALKEDCIAYICREILRLGLAHLHAHVHRIKQNVLLTENAELVDFGVSQAQLDR 180
Db 121 KGNALKEDCIAYICREILRLGLAHLHAHVHRIKQNVLLTENAELVDFGVSQAQLDR 180

Qy 181 TVGRNTTGTTPYMAPEVIACDENPDATYDYSIDISLIGTATAIEMAGAPPLCDMHPMR 240
Db 181 TVGRNTTGTTPYMAPEVIACDENPDATYDYSIDISLIGTATAIEMAGAPPLCDMHPMR 240

Qy 241 ALFLTPRNPPLKSKWKKPFIDTCLIKTYLSRPTQLKFPPIRQDPTQROVRI 300
Db 241 ALFLTPRNPPLKSKWKKPFIDTCLIKTYLSRPTQLKFPPIRQDPTQROVRI 300

Qy 301 QLKDHIDRSKRKGKEETEYEGSSEEDDSHGEGPSSIMNVPGESTLRLREFLRQQ 360
Db 301 QLKDHIDRSKRKGKEETEYEGSSEEDDSHGEGPSSIMNVPGESTLRLREFLRQQ 360

Qy 361 ENKNSSEALKQQQQQLQQQQQDPEAHIKHLHQRORRIEEOKEERRRVEEQORRERQK 420
Db 361 ENKNSSEALKQQQQQLQQQQQDPEAHIKHLHQRORRIEEOKEERRRVEEQORRERQK 420

Qy 421 LOEKEQQRLEDMQALREBERROAERQ-----EY 451
Db 421 LOEKEQQRLEDMQALREBERROAERQEQEYIRHLEBEQRQLETLQQQLQEQALLLEY 480

Qy 452 KRKQLEORQSRLOQEQEHAYLKSQQQQQQQQLQKQQQQQLLPGDRKPLVHYGRGM 511
Db 481 KRKQLEORQSRLOQEQEHAYLKSQQQQQQQQLQKQQQQQLLPGDRKPLVHYGRGM 540

Qy 512 NPADKPAAWEVEERTRMNKQNSPLAKSGSTGPEPPIFOASGPGPGPLSQTPPMQRP 571
Db 541 NPADKPAAWEVVAHR-----VELKPYAAPVPRS----- 569

Qy 572 VEPQGPBKSLQDQTRNLAAFPASHDDPDAIPATATPSARGAVIRONSPTSGGPGS 631
Db 570 -----QSLQDQTRNLAAFPASHDDPDAIPATATPSARGAVIRONSPTSGGPGS 621

Qy 632 PNPPAWVPDNEAPPKVQRTSSIALNTSAGGSRPAQAVRARPRNSAWQIVLQRA 691
Db 622 PNPPAWVPDNEAPPKVQRTSSIALNTSAGGSRPAQAVRA----- 665
```

```
Qy 692 ERGTPKPPGPPAQPFPNASSNDLRRSDPGWERSDVLPSHGHLPQAGSLERNRYGA 751
Db 666 -----SNPDLRRSDPGWERSDVLPSHGHLPQAGSLERNRYGV 704

Qy 752 SSKLDSSPVLSPGNKAKPDDHRSRPRPA-----DFVLLKERTLDDEAPRPKAMDY 803
Db 705 SSKPDSPPVLSPGNKAKPDDHRSRPRPASVKRAIGEDFVLLKERTLDDEAPRPKAMDY 764

Qy 804 SSSSEEVESSEDDDEEGEGGPAEGSRDTPGGRSDGTDSTVSMVHVHVEETGTQPPYGG 863
Db 765 SSSSEEVESSEDDDEEGEGGPAEGSRDTPGGRSDGTDSTVSMVHVHVEETGTQPPYGG 824

Qy 864 GTMVVQRTPEERNLLHADNSGYNLNDVVPQSPSPSTENSQGSPSPKSGSDVQSRLV 923
Db 825 GTMVVQRTPEERNLLHADNSGYNLNDVVPQSPSPSTENSQGSPSPKSGSDVQSRLV 884

Qy 924 KAPKSSFTMFVDLGIYQPGSGSDSIPTALVGGEGTRLDQLQYDVRKGSVVNVNPTNTR 983
Db 885 KAPKSSFTMFVDLGIYQPGSGSDSIPTALVGGEGTRLDQLQYDVRKGSVVNVNPTNTR 944

Qy 984 AHSTPEIRKYKRFNSEILCAALWGNVLLVGTENGMLLDRSGGKGYGLIGRRRFOOM 1043
Db 945 AHSTPEIRKYKRFNSEILCAALWGNVLLVGTENGMLLDRSGGKGYGLIGRRRFOOM 1004

Qy 1044 DVLEGLNLLITISGRNKLRYVYLSWLRNKLHNDPEVEKKQGTTVGDMEGCGHYRVVK 1103
Db 1005 DVLEGLNLLITISGRNKLRYVYLSWLRNKLHNDPEVEKKQGTTVGDMEGCGHYRVVK 1064

Qy 1104 YERIKFLVIALKSSVEVYAWAPKPYHKFMFKSPADLPHRPLLDVLTVEEQRLKVIYGS 1163
Db 1065 YERIKFLVIALKSSVEVYAWAPKPYHKFMFKSPADLPHRPLLDVLTVEEQRLKVIYGS 1124

Qy 1164 SAGHAVDVDSGNSYDIYIPVHIQSITPHAIIFLPNTDGNEMLLCYEDEGVYVNTYGR 1223
Db 1125 SAGHAVDVDSGNSYDIYIPVHIQSITPHAIIFLPNTDGNEMLLCYEDEGVYVNTYGR 1184

Qy 1224 IKDVVLQWGMPTSVAVICSNQIMGWGKAEIRSVETGHLDDGVFMHKAORLKLFCERN 1283
Db 1185 IKDVVLQWGMPTSVAVICSNQIMGWGKAEIRSVETGHLDDGVFMHKAORLKLFCERN 1244

Qy 1284 DKVFFASVRSGSSQVYFMTLNRNCINMW 1312
Db 1245 DKVFFASVRSGSSQVYFMTLNRNCINMW 1273

RESULT 15
US-09-789-390-65
; Sequence 65, Application US/09/789390
; Publication No. US20030059768A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: MacDougall, John
; APPLICANT: Shimkets, Richard A
; APPLICANT: Spaderna, Steven K
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-692
; CURRENT APPLICATION NUMBER: US/09/789,390
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/185,548
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/199,957
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/184,951
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 60/185,967
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/197,723
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 65
; LENGTH: 1303
```



```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(1303)
; OTHER INFORMATION: Wherein Xaa is any amino acid as defined in the
; OTHER INFORMATION: specification.
US-09-789-390-65

Query Match      78.3%; Score 5422.5; DB 10; Length 1303;
Best Local Similarity 79.9%; Pred. No. 11e-241;
Matches 1071; Conservative 0; Mismatches 204; Indels 65; Gaps 3;

Qy 1 MGDPAARSLLDIDLSALRDPAGIFELVEVVGNGTYGVYKGRHVKTGQLAAIKVMDVTE 60
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1 MGDPAARSLLDIDLSALRDPAGIFELVEVVGNGTYGVYKGRHVKTGQLAAIKVMDVTE 60
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 61 DEEBEIKQBINMLKKYSHRNRIATYTGAPIKSPGNDQDLWLVMEFCAGSVTLVKNT 120
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 61 XXXXXIKQBINMLKKYSHRNRIATYTGAPIKSPGNDQDLWLVMEFCAGSVTLVKNT 120
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 121 KGNALKEDCIACIYICREILGLAHLHAHKVVIHRDIKQNVLLTENAENVKLVDFGSAQLDR 180
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 121 KGNALKEDCIACIYICREILGLAHLHAHKVVIHRDIKQNVLLTENAENVKLVDFGSAQLDR 180
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 181 TVGRNTFTGTPTWMAPEVIAIDENPDATYDYSRDIWSLIGITAIEMAEAGAPLCLMDHPWR 240
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 241 ALFLIPRNPPLKSKKSKKFIIDTCLIKTYLSRPTEQLLKFPFIQDPTQVRRI 300
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 241 ALFLIPRNPPLKSKKSKKFIIDTCLIKTYLSRPTEQLLKFPFIQDPTQVRRI 300
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 301 QLKDHIDRSRKKEKEETEYEVYSGSEEDDSDGEGEPSSIMVPGESTLRREFLRQQ 360
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 301 QLKDHIXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXPSIMVPGESTLRREFLRQQ 360
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 361 ENKNSSEALKOQQOQLQOQOORDEPAHIKHLHQRRIEQKEERRRVEEQORREREQRK 420
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 361 ENKNSSEALKXXXXXXXXXXDRDEPAHIKHLHXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 420
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 421 LOEKQQRLEDMQALRREERQREBOEYKQLEQROSERLQRLQOEHAYLSLQ 480
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 421 XXXXXXXXXXXXDMQALXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXAYLSLQ 480
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 481 QOQQOQLQOQOQLLPGDKRPLHYGRGMNADKPAWAREVEERTMKNQOONSPLAKS 540
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 481 XXXXXXXXXXXXXXXXPGDKRPLHYGRGMNADKPAWAREVEERTMKNQOONSPLAKS 540
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 541 KPGSTGPEPPIPOASFGPPGLSQTTPMQRPVPEQEGPHK----- 580
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 541 KPGSTXXXXXXXXXXXXXXXXXXXXXXXXXXXXXMQRPVPEQEGPHKSIVAHRVPLKPYAAPVPSQ 600
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 581 SLQDQPTRNLAAFPASHDDPAIPATPATPSARGAVIRQNSDPTSEGGPGSPNPAAVVRP 640
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 601 SLQDQPTRNLAAFPASHXXXXXXXXXXXXXXXXXXXXXRGAVIRQNSDPTSEGGPGSPNPAAVVRP 660
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 641 DNEAPPKVPORSTSSIALNTALNTSGAGSRPAQAVRARPRNSAWQIYLQRAERGTPKPPG 700
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 661 DNEAPPKVPORSTSSIALNTALNTSGAGSRPAQAVRA----- 695
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 701 PPAQPPGPPNASSNPDRLRRSDFGWERSDVLPAASHGLHPQAGSLERNRVGASSKLDSPV 760
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 696 -----SNPDRLRRSDFGWERSDVLPAASHGLHPQAGSLERNRVGSSKPDSPV 743
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 761 LSPGNKAKPDDHRSRGRPA-----DFVLLKERTLDEAPRPPPKKAMDYSSSEEVES 812
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 744 LSPGNKAKPDDHRSRGRPASRYKRAIGEDFVLLKERTLDEAPRPPPKKAMDYXXXXXXXXXX 803
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 813 SEDDDEEGGPAEGSRDTPGGRSDGTDTSVSTMVVDHVEEITGTQPPYGGTMMVQRTTP 872
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy 804 XXXXXXXXXXXXXXXXRRUTPGGRSDGTDTSVSTMVVDHVEEITGTQPPYGGTMMVQRTTP 863
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: August 28, 2004, 01:03:28
Job time : 88 sec